REFERENCES

Banach, S. (1932). *Théorie des opérations linéares*, Warsaw.

Burkill, J. C. (1932a). The approximately continuous Perron integral, *Math. Z.*, **34**, 270–78.

Burkill, J. C. (1932b). The Cesàro–Perron integral, *Proc. Lond. math. Soc.*, (2)**34**, 314–22.

Burkill, J. C. (1935). The Cesàro–Perron scale of integration, *Proc. Lond. math. Soc.*, (2)**39**, 541–2.

Burkill, J. C. (1936). Fractional orders of integrability, *J. Lond. math. Soc.*, **11**, 220–26.

Burkill, J. C. (1951). Integrals and trigonometric series, *Proc. Lond. math. Soc.*, (3)**1**, 46–57.

Cameron, R. H. and Martin, W. T. (1941). An unsymmetrical Fubini theorem, *Bull. Amer. math. Soc.*, **47**, 121–5.

Getchell, B. C. (1935). On the equivalence of two methods of defining Stieltjes integrals, *Bull. Amer. math. Soc.* (2)**41**, 413–8.

Henstock, R. (1946). On interval functions and their integrals, *J. Lond. math. Soc.* **21**, 204–9.

Henstock, R. (1948). On interval functions and their integrals II, *J. Lond. math. Soc.* **23**, 118–23.

Henstock, R. (1960a). A new descriptive definition of the Ward integral, *J. Lond. math. Soc.* **35**, 43–8.

Henstock, R. (1960b). The use of convergence factors in Ward integration, *Proc. Lond. math. Soc.* (3)**10**, 107–21.

Henstock, R. (1960c). The equivalence of generalized forms of the Ward, variational, Denjoy–Stieltjes and Perron–Stieltjes integrals, *Proc. Lond. math. Soc.* (3)**10**, 281–303.

Henstock, R. (1961a). N-variation and N-variational integrals of set functions, *Proc. Lond. math. Soc.* (3)**11**, 109–33.

Henstock, R. (1961b). Definitions of Riemann type of the variational integrals, *Proc. Lond. math. Soc.* (3)**11**, 402–18.

Hildebrandt, T. H. (1926). The Borel theorem, *Bull. Amer. math. Soc.* **32**, 423–74.

Jeffery, R. L. and Miller, D. S. (1945). Convergence factors for generalized integrals, *Duke math. J.* **12**, 127–42.

Jeffery, R. L. (1951). The theory of functions of a real variable, *Math. Expos.* No. 6, Univ. of Toronto Press.

THEORY
OF
INTEGRATION

RALPH HENSTOCK, M.A., PH.D.
Reader in Pure Mathematics
The Queen's University, Belfast

LONDON
BUTTERWORTHS
1962

ENGLAND BUTTERWORTH & CO. (PUBLISHERS) LTD.
London 88 Kingsway, W.C.2

AFRICA BUTTERWORTH & CO. (AFRICA) LTD.
DURBAN 33/35 Beach Grove

AUSTRALIA BUTTERWORTH & CO. (AUSTRALIA) LTD.
SYDNEY 6-8 O'Connell Street
MELBOURNE 430 Bourke Street
BRISBANE 240 Queen Street

CANADA BUTTERWORTH & CO. (CANADA) LTD.
TORONTO 1367 Danforth Avenue, 6

NEW ZEALAND BUTTERWORTH & CO. (NEW ZEALAND) LTD.
WELLINGTON 49/51 Ballance Street
AUCKLAND 35 High Street

U.S.A. BUTTERWORTH INC.
WASHINGTON, D.C. 7235 Wisconsin Avenue, 14

Suggested U.D.C. number: 517·3

©
Butterworth & Co (Publishers) Ltd.
1962

J. W. ARROWSMITH LTD., BRISTOL 3

PREFACE

It gives me great pleasure to acknowledge the advice of Professor S. Verblunsky, who has read the text, and also the following lectures and notes—

(i) Lectures on Lebesgue integration by J. C. Burkill (1942), R. G. Cooke (1944), A. S. Besicovitch (1947).
(ii) Notes on Stieltjes integration by P. Dienes (1944).
(iii) Lectures on statistics and stochastic processes by J. O. Irwin (1943), J. Wishart (1947), M. S. Bartlett (1947).

Most of the book has been given in lecture form at the Queen's University, Belfast. Two research students, Mr. James McGrotty and Mr. Chan Kai Meng, have read the proofs. Nevertheless, I am solely responsible for errors and omissions that remain.

THE QUEEN'S UNIVERSITY
BELFAST

RALPH HENSTOCK

CONTENTS

CHAPTER PAGE

Preface v

1. HISTORICAL INTRODUCTION . . . 1
1. General introduction—2. Areas, and the differential and integral calculus—3. Riemann, Riemann–Stieltjes and Burkill integration—4. A new approach to Newton's and Riemann's integrals—5. The Pollard–Getchell integral—6. Lebesgue and Radon integration—7. Special and general Denjoy integrals—8. Perron, Ward, and Burkill's approximate Perron integrals—9. Extensions to 1958—10. The N-, N-variational, and Riemann-complete integrals

2. THE RIEMANN-COMPLETE INTEGRAL . 13
11. Notation—12. Countability—13. The real line—14. Open sets—15. Limit-points and closed sets—16. Left-complete and right-complete families—17. The integration of point functions—18. Functions of intervals—19. Distributive functional—20. Order relation—21. Additivity in elementary sets—22. Two integrability results—23. Stieltjes transformation

3. VARIATIONAL PROPERTIES OF THE INTEGRAL 39
24. The variational integral—25. Order relation and integrability—26. Variation—27. Inner variation—28. Properties of the variation and inner variation—29. Functions of generalized bounded variation—30. The variation of variationally equivalent functions—31. Special results—32. Sierpinski's lemma and the inner variation—33. Integration by parts

4. DIFFERENTIATION 74
34. Differentiation of VB^* and VBG^* functions—35. The differentiation of variational integrals

5. LIMITS UNDER THE INTEGRAL SIGN . 82
36. Monotone sequences and functions—37. Majorized sequences and functions—38. Absolute integration—39. Limits of step functions—40. An absolute integral is AC^*, and the integrability of a product

6. DOUBLE INTEGRALS AND FUBINI'S THEOREM 100
41. The plane—42. Functions of rectangles—43. The integration of rectangle functions—44. Fubini's theorem—45. Further results

7. THE CAUCHY AND DENJOY EXTENSIONS, AND THE INTEGRAL IN AN INFINITE RANGE 113
46. The Cauchy extension—47. The integral in an infinite range—48. The Denjoy extension

8. CONNECTIONS WITH EARLIER INTEGRALS 121
49. The integrals of Sections 2 to 7—50. The Perron and Ward integrals

9. LINEAR TOPOLOGICAL SPACES, YOUNG'S INEQUALITY AND INTEGRATION . . 127
51. Linear spaces—52. Linear topological spaces—53. Young's inequality—54. Mean convergence—55. The complete normed linear space $L^p(p > 1)$—56. The Riemann-complete integral for functions with values in a linear topological space—57. Functions of bounded variation

10. INTEGRATION IN STATISTICS . . . 148
58. Introduction—59. Classification and relative frequency—60. Probability—61. Probability in three areas—62. Independence—63. The Neyman-Pearson theory of tests of simple hypotheses—64. Generating functions, moments

References 162

Index 165

The Newton & Riemann ∫s.

HISTORICAL INTRODUCTION

1. General introduction

Most students begin real variable theory with the ~~differential~~ calculus, ~~proceeding to the~~ defining integrals as the inverse of ~~the derivative~~ differentiation. In essentials ~~they~~ using the ideas of Newton and Leibnitz. Some never go beyond this stage; much is done in physics and applied mathematics by using ~~this~~ the calculus alone, without recourse to more modern ideas. However, it is becoming increasingly clear to scientists in general, ~~and particularly to~~ especially to those interested in statistics, statistical physics ~~and~~ or quantum theory, that such a limitation of the pure mathematics [illegible] is a restriction, and that the use of a better integration ~~tools~~ process will result in a greater ease of application. The mathematician or scientist who wishes to go beyond the Newtonian inverse of a derivative could advantageously take for his tools the two new and equivalent integrals of this book. He will find that the usual theorems follow. In particular,

(i) the integral is additive in the functions to be integrated, and in the ~~intervals in~~ n-dim rectangles over which they are integrated (Theorem 19.1, p. 27; Theorem 21.1, p. 31);

(ii) if the convergent sequence $\{f_n(x)\}$ is bounded by M independent of n and x, and if each $f_n(x)$ is integrable, then

$$\int_a^b \lim_{n\to\infty} f_n(x)\,\mathrm{d}x = \lim_{n\to\infty}\int_a^b f_n(x)\,\mathrm{d}x$$

(special case of Theorem 37.1, p. 85);

(iii) if in (ii), $f_n(x)$ is the sum of $g_m(x)$ for $m = 1, 2, \ldots, n$, then

$$\int_a^b \sum_{m=1}^{\infty} g_m(x)\,\mathrm{d}x = \sum_{m=1}^{\infty}\int_a^b g_m(x)\,\mathrm{d}x$$

(iv) $\dfrac{\mathrm{d}}{\mathrm{d}x}\displaystyle\int_a^x f(t)\,\mathrm{d}t = f(x)$

at almost all points x (special case of Theorem 35.1, p. 78);

(v) if, for each y, Y, $f(x, y)$ is integrable and

$$|f(x, Y)-f(x,y)| \leqslant M|Y-y|,$$

where M is independent of x, y, Y, then

$$\frac{\mathrm{d}}{\mathrm{d}y}\int_a^b f(x, y)\,\mathrm{d}x = \int_a^b \frac{\partial f(x, y)}{\partial y}\,\mathrm{d}x$$

(special case of Theorem 37.3, p. 88).

The theory is developed from the beginning, and only a knowledge of real and complex numbers is assumed. However, a familiarity with the ϵ, δ technique would be an advantage.

The ordinary student is advised to omit all sections that are starred, and also Chapter 8. These topics are intended for the more mature student who is acquainted with the Lebesgue integral at least. To him it may be pointed out that the very elegance and apparent finality of Lebesgue theory has caused mathematical inventiveness to move to other fields. Because of the Riesz representation theorem the integral has been regarded as nothing more than a linear functional. But integration is far more than this, as the present book partially shows. Further, there is no longer any necessity for considering separately the Riemann, Riemann–Stieltjes, Burkill, Moore-Pollard–~~Getchell~~, Lebesgue, Radon, special Denjoy, Perron and Ward integrals, for all these are included in the Riemann-complete and variational integrals of this book. Measure theory is not needed at the beginning, and the measure properties that are required later are obtained in a far simpler way than usual. Contrary to the general view, complete additivity of the measure is *not* required in order that we should have theorems of Lebesgue type; here we find that finite additivity is enough. In particular, Fatou's lemma is still valid, so that the Riesz–Fischer theorem and the completeness of L^p spaces follow as usual. For the latter see Chapter 9, in which we extend the integral to functions with values in linear topological spaces. We omit all integrals using convergence factors, and for these we refer to the literature (Henstock, (1960b, c; 1961a, b)). For simplicity we also restrict the functions to be functions of points or intervals on the real axis, or of points or rectangles on the plane. The n-dimensional case is a straightforward generalization, and the extension to an abstract space is given in Henstock (1961a, b).

But it is not advisable to restrict ourselves entirely to pure mathematics. There is ample precedent for a pure mathematician to consider in detail some applications to statistics (Chapter 10). This chapter is bound to be inconclusive and to leave much unsaid because of limitation of space, so that I have made an arbitrary selection of topics that illustrate the main theory. The Central Limit and allied theorems are adequately covered in other books, and so are omitted.

The first chapter is divided into two, because of the requirements of two kinds of reader. The beginner can read Sections 2 to 4 and then proceed to Chapter 2. The remaining sections are for the notice of the integration student who has already dealt with Lebesgue integration at least, to explain the relations between the new integrals and the old. We give a very brief summary of the position of integration theory up to, say, 1958. Because of the brevity many illustrious names have to be omitted, and we concentrate on generalizations suitable for problems in trigonometric series.

Blaise Pascal 1623-1662, David Hilbert 1862-1943

2. Areas, and the differential and integral calculus

In the simplest case the process of integration is the adding together of areas of non-overlapping elementary figures, and then the taking of some kind of a limit. The Greeks computed many simple areas, and then ~~the~~ methods were ~~being~~ systematized through the years, ~~and~~ culminating in the *method of exhaustions* of Eudoxus (*c*.408–355 B.C.) and Archimedes (*c*.287–212 B.C.). This ~~method~~ was the first crude limit process; ~~and~~ they used the geometry of the figures to fit a sequence of non-overlapping triangles inside each main figure that finally exhausts the area. By this means they found the areas of the circle and sections of parabolas, for example, but could not define a general non-negative polynomial, and so could not compute the area under its curve.

The second approach to integration lies in inverting the result of differentiating a known function. The operation of differentiation was first systematized by I. Newton (1642–1727) and G. W. Leibnitz (1646–1716). To each of a certain class of functions f for which the derivative $Df = \mathrm{d}f/\mathrm{d}x$ exists, say, for x in $a \leqslant x \leqslant b$, we make correspond that derivative, so that we can regard D as an *operator*. It obeys the following rules. If f, g are differentiable functions of x in $a \leqslant x \leqslant b$, and if α, β are constants, then in $a \leqslant x \leqslant b$ we have

$$D(\alpha f+\beta g) = \alpha Df+\beta Dg \tag{2.1}$$

$$D(fg) = (Df)g+f(Dg) \tag{2.2}$$

$$D\{f(g(x))\} = (\mathrm{d}f/\mathrm{d}g)Dg \tag{2.3}$$

If g and f = h/g are diff:

$$D\alpha = 0 \tag{2.4}$$

The rule for division is obtained from (2.2); if g, h, and $f = h/g$ are diff. then

$$Dh = (Df)g+f(Dg)$$

$$D(h/g) = Df = \{Dh-(h/g)Dg\}/g$$

A function H of points x *is an indefinite Newton integral* of a known finite function f in $a \leqslant x \leqslant b$, if $DH = f$ in that interval. The functions that Newton integrated are all continuous, but we can ignore that limitation. Then the *definite Newton integral in $a \leqslant x \leqslant b$ is $H(b)-H(a)$*. We can write H as

$$H = D^{-1}f = (NL)\int f\,\mathrm{d}x,\ H(b)-H(a) = (NL)\int_a^b f\,\mathrm{d}x$$

where NL stands for Newton–Leibnitz. This definition of the integral is *descriptive*. No method of construction is offered, but we are given its properties so that we can recognize it if it is ~~produced~~ constructed in ~~another~~ some way. Because of this we have to prove that if H and H_1 are both indefinite Newton integrals of the same function f in $a \leqslant x \leqslant b$, then

$$H(b)-H(a) = H_1(b)-H_1(a) \tag{2.5}$$

To prove (2.5) we note that by (2.1)

$$D(H-H_1) = f-f = 0$$

so that in particular $H-H_1$ is continuous, and then the mean value theorem gives (2.5). [illegible]

From (2.1) we obtain the distributivity of the Newton integral, namely,

$$D^{-1}(\alpha f+\beta g) = \alpha D^{-1}f+\beta D^{-1}g \tag{2.6}$$

From (2.2; 2.6) we have the formula for *integration by parts*,

$$D^{-1}(gDf)+D^{-1}(fDg) = fg$$

$$(NL)\int\left(\frac{f\,\mathrm{d}g}{\mathrm{d}x}\right)\mathrm{d}x = fg-(NL)\int\left(\frac{g\,\mathrm{d}f}{\mathrm{d}x}\right)\mathrm{d}x \tag{2.7}$$

From (2.3) we have

$$f(g(x)) = (NL)\int\frac{\mathrm{d}f}{\mathrm{d}g}\cdot\frac{\mathrm{d}g}{\mathrm{d}x}\,\mathrm{d}x$$

and replacing $\mathrm{d}f/\mathrm{d}g$ by f_1,

$$(NL)\int f_1(g)\,\mathrm{d}g = (NL)\int f_1\cdot\frac{\mathrm{d}g}{\mathrm{d}x}\,\mathrm{d}x \tag{2.8}$$

the formula for *integration by substitution*.

When we have defined more general integrals we will see that the formulae (2.5; 2.6; 2.7; 2.8) are in some sense still true for them.

The integration of a polynomial in x is now easy, but some simple functions cannot be integrated. It can be proved that if DH exists in $a \leqslant x \leqslant b$, and if γ is a number between $H'(a)$ and $H'(b)$, then there is a ξ in $a \leqslant \xi \leqslant b$ such that $H'(\xi) = \gamma$. It follows that if f is zero for x less than $\frac{1}{2}(a+b)$, and is 1 otherwise, then f does not have a Newton integral in $a \leqslant x \leqslant b$.

3. Riemann, Riemann–Stieltjes and Burkill integration

G. F. B. Riemann (1826–66) gave in 1854 the following definition of the definite integral of a function f in $a \leqslant x \leqslant b$. Let

$$a = x_0 < x_1 < \dots < x_n = b \tag{3.1}$$

be a division of $a \leqslant x \leqslant b$ into smaller intervals, let ξ_j be a point of the interval $x_{j-1} \leqslant x \leqslant x_j$, and consider the sum

$$S = \sum_{j=1}^{n} f(\xi_j)(x_j-x_{j-1}) \tag{3.2}$$

The number *I is the definite Riemann integral of f in $a \leqslant x \leqslant b$*, if to each $\epsilon > 0$ there is a $\delta > 0$ such that

$$|S-I| < \epsilon \tag{3.3}$$

whenever

$$x_{j-1} \leqslant \xi_j \leqslant x_j < x_{j-1}+\delta \quad (j = 1, 2, \dots, n) \tag{3.4}$$

J. G. Darboux (1842–1917) made the following modification when f is real. He replaced $f(\xi_j)$ by the *supremum* (*least upper bound*) *of f in $x_{j-1} \leqslant x \leqslant x_j$*, and obtained an *upper sum*. For a *lower sum* he replaced $f(\xi_j)$ by the *infimum*

"Mémoire sur les fonctions discontinues" (Ann. Éc. Norm., 1875) *(greatest lower bound) of f in $x_{j-1} \leqslant x \leqslant x_j$.* If f is non-negative, with a given graph, and if we take a division (3.1) of $a \leqslant x \leqslant b$, then the upper Darboux sum is the sum of the areas of rectangles with bases the intervals $x_{j-1} \leqslant x \leqslant x_j$, and with just sufficient height to include the graph. The lower Darboux sum is the sum of the areas of rectangles with the same bases, but lying just below the graph. When f is real it is clear that for suitable choice of the ξ_j, the S of (3.2) can be taken arbitrarily near to the upper sum, and for another choice, arbitrarily near to the lower sum, so that the Darboux modification does not alter the Riemann integral of a real function. Thus if a real function has a Riemann integral in $a \leqslant x \leqslant b$ it must be bounded there. From this we can show that not every Newton integral is a Riemann integral. For

$$H(x) = x^2.\sin(1/x^2)\ (x \neq 0),\ H(0) = 0 \tag{3.5}$$

is differentiable everywhere, the derivative being unbounded in the neighbourhood of $x = 0$. However, not every Riemann integral is a Newton integral, for the Riemann integral of the last function of Section 2 exists in $a \leqslant x \leqslant b$, and is equal to $\frac{1}{2}(b-a)$. There is a common region, for the Riemann and Newton integrals of a continuous function exist and are equal. The Riemann integral cannot integrate every bounded function, for if

$$f(x) = \begin{cases} 1 \ (x \textit{ rational}) \\ 0 \ (x \textit{ irrational}) \end{cases} \tag{3.6}$$

then any upper Darboux sum is $b-a$, while any lower Darboux sum is 0. Thus f does not have a Riemann integral (nor a Newton integral).

The Riemann method has been modified in many ways. T. J. Stieltjes (1856–94) used another function g, replacing $x_j - x_{j-1}$ in (3.2) by

Recherches sur les fractions continues (Ann. de la Fac. des Sc. de Toulouse 8 (1894) 1-122.

$$g(x_j) - g(x_{j-1})$$

The resulting integral is now called the *Riemann–Stieltjes integral*. J. C. Burkill replaced $f(\xi_j)(x_j - x_{j-1})$ in (3.2) by a function $h(x_{j-1}, x_j)$ of the interval from x_{j-1} to x_j, obtaining the *Burkill integral*. It is clear that the Riemann, Riemann–Stieltjes and Burkill integrals are *constructive* when they exist. For we can construct the definite integral I by finding the limit of sums (3.2) for Riemann integrals, or the corresponding sums for the other integrals, for special sequences of divisions (3.1) and special choices of the ξ_j, e.g. we can take $\xi_j = x_j = a + (b-a)j/2^n$ for $j = 0, 1, 2, \ldots, 2^n$ and $n = 1, 2, \ldots$

4. A new approach to Newton's and Riemann's integrals

Newton's definite integral can be written as a simple example of the kinds of integrals that we wish to study. Let H be the Newton indefinite integral of a finite function f in $a \leqslant x \leqslant b$. Then $DH = f$ there. Thus, given $\epsilon > 0$, there is a $\delta > 0$, depending on ϵ and x, such that for $0 < |t-x| \leqslant \delta$,

$$\left|\frac{H(t)-H(x)}{t-x} - f(x)\right| \leqslant \epsilon,$$

$$|H(t) - H(x) - f(x)(t-x)| \leqslant \epsilon|t-x| \tag{4.1}$$

Here, all intervals from t to x, where either $x - \delta \leqslant t < x$ or $x < t \leqslant x + \delta$, have this property. Suppose that from these intervals we can construct a division (3.1) of $a \leqslant x \leqslant b$. Then for ξ_j equal to one or other of x_{j-1}, x_j, we use (4.1) and obtain

$$\begin{aligned} &|H(b) - H(a) - \sum_{j=1}^{n} f(\xi_j)(x_j - x_{j-1})| \\ &\quad = |\sum_{j=1}^{n} \{H(x_j) - H(x_{j-1}) - f(\xi_j)(x_j - x_{j-1})\}| \\ &\quad \leqslant \sum_{j=1}^{n} |H(x_j) - H(x_{j-1}) - f(\xi_j)(x_j - x_{j-1})| \\ &\quad \leqslant \sum_{j=1}^{n} \epsilon(x_j - x_{j-1}) = \epsilon(b-a) \qquad \cancel{(3.2)} \end{aligned}$$

showing that for the special ξ_j and special divisions, the sum (3.2) tends to $H(b) - H(a)$ as $\epsilon \to 0$. Thus if the construction of divisions is possible, Newton's integral has a definition as an integral of Riemann type. In Chapter 2 we show that the construction is in fact possible. Further, the interval functions $f(\xi_j)(x_j - x_{j-1})$ are of two types,

$$f(x_{j-1})(x_j - x_{j-1}), \quad f(x_j)(x_j - x_{j-1}) \tag{4.2}$$

right-hand and left-hand, where the ξ_j is regarded as fixed, while the other end of the interval lies in a certain neighbourhood of ξ_j.

Again, in Riemann's definition we use interval functions

$$f(\xi_j)(x_j - x_{j-1})(x_{j-1} \leqslant \xi_j \leqslant x_j) \tag{4.2}$$

If $\xi_j = x_{j-1}$ or x_j, we again obtain an interval function (4.2). If $x_{j-1} < \xi_j < x_j$ we put

$$f(\xi_j)(x_j - x_{j-1}) = f(\xi_j)(\xi_j - x_{j-1}) + f(\xi_j)(x_j - \xi_j)$$

a sum of the two kinds in (4.2). Similarly for the Riemann–Stieltjes integral. There is no ξ_j in Burkill integration, so that it does not matter whether an interval is counted as left-hand or right-hand, the h is the same.

This section gives us two ideas that are developed in Chapter 2; namely, the use of intervals from t to x, for $|t-x| \leqslant \delta(x)$, and the use of two interval functions, one 'left-hand' and one 'right-hand'. The beginner can now proceed to Chapter 2.

The rest of Chapter 1 is designed for the student of integration who wishes to connect his previous knowledge with the theory of this book.

*5. The Moore-Pollard~~-Getchell~~ integral

In Riemann–Stieltjes integration S. Pollard (1894–1945) and B. C. Getchell modified (3.4), supposing that for each $\epsilon > 0$, there is a division

$$a = x_0' < x_1' < \ldots < x_p' = b \tag{5.1}$$

with the property that if (3.1) is a subdivision of (5.1) and $x_{j-1} \leqslant \xi_j \leqslant x_j$ for $j = 1, 2, \ldots, n$, then (3.3) is true. In this case *I is the Pollard-Getchell integral of f*. See Pollard (1923), Getchell (1935).

Every Riemann–Stieltjes integral is a Pollard–Getchell integral, for we can ensure (3.4) by taking, in (5.1)

$$x'_k - x'_{k-1} < \delta \quad (k = 1, 2, \ldots, p)$$

The Pollard–Getchell integral enables us sometimes to integrate f with respect to g when f and g have common discontinuities and the Riemann–Stieltjes integral does not exist. The Pollard–Getchell modification can be applied to the Burkill integral, and a suitable construction of the divisions (5.1) in this case are given in Henstock (1946 and 1948). Clearly the Pollard–Getchell integral is constructive if suitable divisions (5.1) can be constructed.

The integrals of Sections 3 and 5 are of Riemann type and share most of the limitations of the Riemann integral. In particular, they cannot integrate the limits of all bounded convergent sequences of integrable functions, nor can they integrate all derivatives, so that more drastic modifications are necessary, such as will be given later in this book when we define the Riemann-complete integral.

*6. Lebesgue and Radon integration

H. Lebesgue (1875–1941) made a great step forward by defining first the measure of a set of real points, and secondly, his integral. J. Radon (1887-195 replaced Lebesgue measure by a measure suitable for Stieltjes integrals. Lebesgue type integrals will be defined in Chapter 8. Other definitions have been given by W. H. Young (1863–1942) and P. J. Daniell (1889–1946).

If f is Riemann-integrable then it is Lebesgue-integrable. But the converse is not always true, as the Lebesgue integral of (3.6) exists and is equal to 0. Further, if a sequence of Lebesgue-integrable functions is bounded and convergent then the limit-function is also Lebesgue-integrable, the integral being equal to the limit of the integrals of members of the sequence. If f is Lebesgue-integrable then so is $|f|$, so that we have an *absolute integration*. This last result can be used to show that the derivative of (3.5) cannot be Lebesgue-integrable, so that Newton's integral is not included in Lebesgue's.

The situation is radically different when we pass to Stieltjes-type integrals. If a Radon measure is obtained from a point function g, then g must be of bounded variation, whereas, for example, the Riemann–Stieltjes integral of 1 with respect to g in $a \leqslant x \leqslant b$ always exists and is equal to $g(b) - g(a)$, even if g is not of bounded variation. If g is of bounded variation and if the Riemann–Stieltjes integral of f with respect to g exists, so does the corresponding Radon integral. The two integrals may not be equal unless g is also continuous. The Radon integral can also integrate the limit-function of a bounded sequence of Radon-integrable functions, if the integrator g remains fixed throughout. We obtain the Lebesgue and Radon integrals in a set M contained in an interval I, if we multiply the integrand by the characteristic function of M and integrate in I. An unbounded M needs an infinite interval. Thus we need only integrate in intervals. Clearly also, the integrals are constructive.

W. H. Young "On the integration with respect to a function of bounded variation"
PLMS(1913)
J. Radon "Theorie und Anwendungen der absolut additiven Mengenfunktionen"

Sitzungsberichte der Kais. Ak. d. Wiss. in Wien, Bd 122 Abt IIa (1913)
28-33.

*7. Special and general Denjoy integrals

By using two methods of extension of the Lebesgue integral, A. Denjoy defined his *special Denjoy integral*. The two methods are given in Chapter 7, and the special Denjoy integral in Chapter 8. The Lebesgue and Newton integrals are contained in the special Denjoy integral, which in turn is contained in the *general Denjoy integral*, obtained by a modification of the second extension of the Lebesgue integral. But the general Denjoy integral cannot integrate all sums of convergent trigonometric series.

In passing from the Lebesgue to the Denjoy integrals we lose the absoluteness of the integral. Thus $|f|$ need not be Denjoy-integrable even if f is, and the Denjoy integrals are *non-absolute*.

I have proved, Henstock (1960c), that a special Denjoy–Stieltjes integral can be defined if the integrator is fairly smooth, and in particular, if it is continuous and monotone. The proof deals with a more complicated case, but the simpler case is an easy deduction.

The Denjoy-type integrals are again constructive.

*8. Perron, Ward, and Burkill's approximate Perron integrals

O. Perron defined *major and minor functions* of a function f in $a \leqslant x \leqslant b$, and with these defined the *Perron integral*. The simple descriptive definition is given in Chapter 8, and it does not need measure theory. It is equivalent to the special Denjoy integral, in the sense that if one exists then so does the other, and their values are equal. A Perron–Stieltjes integral can be given when g is strictly increasing, but the definition is difficult otherwise. This led A. J. Ward (1936) to define the *Ward integral* as an integral of Perron and Stieltjes type that is again descriptive, and is equivalent to the Perron–Stieltjes and special Denjoy–Stieltjes integrals when they exist. Further, if the Riemann–Stieltjes or Newton integral exists, so does the Ward integral, with the same value.

By a modification of the derivatives used in Perron integration, J. C. Burkill (1932a) has defined an *approximate Perron integral* with approximate derivative equal to the original function.

*9. Extensions to 1958

Many other real variable integrals have been defined, most of them using convergence factors. The first of these were *Burkill's Cesàro–Perron integrals* (Burkill (1932b, 1935 and 1936)), and some of them, and *Burkill's symmetrical Cesàro–Perron integrals* (Burkill (1951)), can integrate all sum-functions of convergent trigonometric series. They cannot integrate all Abel sums of Abel-convergent trigonometric series, and the Abel–Poisson–Perron integral of S. J. Taylor (1956) has been constructed for this purpose. To put the theory of convergence-factor integrals on a more general foundation, R. L. Jeffery and D. S. Miller (1945) have produced general convergence factor integrals of Perron type.

Most real variable integrals can be extended to deal with complex-variable functions, and functions defined in an n-dimensional rectangle or in an abstract space. Integrals have also been defined for functions with values in Banach spaces, or even in linear convex topological spaces.

An approach due to F. Riesz, is to regard an integral as nothing more than a linear functional on a space of functions. This is due to the linearity of the integral, and the Riesz representation theorem, that gives a Stieltjes integration corresponding to each linear functional. But the theory of integration is much wider than that, and the linear functional approach is unduly restrictive.

*10. The N–, N–variational, and Riemann-complete integrals

Since 1958, to simplify the theory of convergence-factor integrals, I combined the approaches of Ward, and Jeffery and Miller, and the result is the *N-integral* of Henstock (1960b). Further theory given in Henstock (1960c and 1961a) make it clear that an equivalent descriptive integral, the *N-variational integral*, can be defined. The simple form of the latter, in which convergence factors are omitted, is the *variational* integral of Henstock (1960a). The next step, in which we travel back to a Riemann-type integral, is the *Riemann-complete integral* of Henstock (1961b). The Riemann-complete and equivalent variational integrals form the subject of this book, and they include Riemann, Riemann–Stieltjes, Pollard–Getchell, Burkill, Lebesgue, special Denjoy, Perron, and Ward integrals, and Radon integrals for a continuous integrator. Differences occur with the Radon integral if the integrator is discontinuous. The Riemann-complete integral is constructive if certain 'complete sets' are constructive, while the variational integral is descriptive.

In Chapter 9 of this book we deal briefly with integrals of functions with values in more general spaces. The subject of the use of convergence factors in integrals is so vast that it would require a separate book to deal justly with the theory. I hope in due course to write such a book.

2

THE RIEMANN-COMPLETE INTEGRAL

11. Notation

As usual in books on integration theory, we begin with notation involving sets and their members. First, an object x may be described in several ways. If x and y represent the same object we write $x = y$. A collection of objects is sometimes called a *set*, and sometimes a *family*. These two terms are regarded as being interchangeable and as meaning the same thing. If an object x is a member of a set M we write $x \in M$. If every member of M is also a member of a set M_1 we write

$$M \subseteq M_1,\ M_1 \supseteq M$$

If M,M_1 have the same members,

$$M \subseteq M_1,\ M_1 \subseteq M$$

and we write $M = M_1$, and regard M and M_1 as being the same set. Otherwise we write $M \neq M_1$ and say that M,M_1 are *distinct*. If

$$M \subseteq M_1,\ M \neq M_1,\ \textit{we write}\ M \subset M_1,\ M_1 \supset M$$

If $\mathscr{M}$ is a family of sets M then

$$\bigcup_{\mathscr{M}} M,\quad \bigcap_{\mathscr{M}} M$$

are the *union and intersection* of sets M of $\mathscr{M}$. The first is the set of all objects each of which lies in *at least one* of the sets of $\mathscr{M}$. The second is the set of all objects each of which lies in *all* sets of $\mathscr{M}$. Usually $\mathscr{M}$ is omitted from the symbol, and there are other slight variations. For instance, if $\mathscr{M}$ is the collection of M_j for $j = 1,2,\ldots$, we write

$$\bigcup_{j=1}^{\infty} M_j,\quad \bigcap_{j=1}^{\infty} M_j$$

Note that we do not need any idea of a limit in these definitions. The union and intersection of two sets M,M_1 are for simplicity written as $M \cup M_1$, $M \cap M_1$. Two sets with no common member, so that the intersection is empty, are called

cup
or
cap

disjoint. If each pair of sets from a family $\mathscr{M}$ of sets is disjoint we say that the sets of $\mathscr{M}$ *are mutually disjoint.* If M is a set of real numbers the *complement* CM *of* M is the set of all real numbers that do not lie in M. Thus $M \cap CM$ is empty while $M \cup CM$ is the set of all real numbers. If $\{M_j\}$ is a sequence of sets then we define

$$\limsup_{j\to\infty} M_j = \bigcap_{n=1}^{\infty}\bigcup_{j=n}^{\infty} M_j, \quad \liminf_{j\to\infty} M_j = \bigcup_{n=1}^{\infty}\bigcap_{j=n}^{\infty} M_j,$$

and if they are the same set we write it as

$$\lim_{j\to\infty} M_j$$

Ex. 11.1. If $M_1 \subseteq M_2 \subseteq \ldots \subseteq M_j \subseteq \ldots$ show that

$$\lim_{j\to\infty} M_j = \bigcup_{j=1}^{\infty} M_j$$

Ex. 11.2. If $M_1 \supseteq M_2 \supseteq \ldots \supseteq M_j \supseteq \ldots$ show that

$$\lim_{j\to\infty} M_j = \bigcap_{j=1}^{\infty} M_j$$

III o.

12. Countability

Suppose that T is a relation between the objects x of a set X and the objects y of a set Y, such that

(a) if x is in X there is a y in Y such that xTy;
(b) if there is a y' *in* Y such that xTy' then $y' = y$;
(c) if y is in Y there is an x in X such that xTy;
(d) if there is an x' in X such that $x'Ty$ then $x' = x$.

In this case the relation T is called a *one-one correspondence,* and we say that *X, Y are similar.* By using this idea it is possible to define the positive integers as *cardinal numbers.* Here we will assume that the real and complex number systems have been defined in some way already, and will only use the idea of similarity to define *countability.*

A set X *is countable* if it is similar to the set of all positive integers. If to each object x we attach its positive integer j and ~~if we~~ use the~~ir~~ natural order of the we obtain the infinite sequence $\{x_j\}$, a notation already used.

Cantor, Crelle's Journal 77 (1874), p. 258

III o.

Theorem 12.1. *A subset* M_1 *of a countable set* M *is finite or countable.*

We can suppose that M_1 is not empty. Then, writing the set M as a sequence, let j_1 be the first integer j such that x_j is in M_1. Let j_2 be the first integer $j > j_1$ such that x_j is in M_1, and so on. If at some stage the process ends, the set M_1 is finite. Otherwise we obtain an infinite subsequence $\{x_{j_n}\}$ of points of M_1. Each point of M_1 is an x_j for some integer j, and so occurs in the subsequence, so that M_1 is countable in this case.

Theorem 12.2. *A finite or countable union of finite or countable sets is finite or countable.*

Let the sets of the union be put in a finite or infinite sequence as $\{X_j\}$, and let the members of X_j be put in a finite or infinite sequence as $\{x_{jk}\}$, for $k = 1,2,\ldots$. We use the diagonal process of G. Cantor (1845–1918), collecting the x_{jk} into groups with equal $j+k$, and then ordering each group by an increasing j. Thus we obtain

$$x_{11}, x_{12}, x_{21}, x_{13}, x_{22}, x_{31}, x_{14}, x_{23}, x_{32}, x_{41}, x_{15}, \ldots$$

We now use this order to consider each x_{jk} in turn, and we discard the x_{jk} if it is equal to an earlier member of the sequence, and we otherwise retain it. The result is a subset S of a finite or countable set, which by Theorem 1 is again finite or countable. Each object of the union is in an X_j and so is an object x_{jk}, which therefore occurs in S. Clearly no other object can occur in S, so that we have proved the theorem.

Cantor, p. 260 Also Jahresbericht der deutschen math. Vereinig 1 (1892) p 75.

III o-

Theorem 12.3. *The set of all real numbers in* $0 \leqslant x < 1$ *is not countable.*

Each x in $0 \leqslant x < 1$ can be written as a decimal, say

$$x = 0{\cdot}x^1x^2\ldots x^n\ldots$$

where x^n takes an integer value from 0 to 9. Some real numbers have a dual representation, for example,

$$0{\cdot}5000\ldots = 0{\cdot}4999\ldots$$

To remove this duality we make the convention that no decimal has the form with $x^n = 9$ for all n greater than some integer.

Suppose that the real numbers between 0 and 1 can be written as a sequence $\{x_j\}$. Then each x_j can be written as

$$x = 0{\cdot}x_j{}^1x_j{}^2\ldots x_j{}^n\ldots$$

and we can define a decimal

$$y = 0{\cdot}y^1y^2\ldots y^n\ldots$$

by the rule

$$y^n = x_n{}^n+1 \ (0 \leqslant x_n{}^n \leqslant 7),\ y^n = 1 \ (x_n{}^n = 8,9)$$

so that $y^n \neq 9$. This decimal differs from each decimal of the sequence, as it differs from x_j in at least the jth decimal place. Also the decimal is a number in $0 \leqslant y < 1$, so that we have a contradiction. This proves the theorem.

The theorem shows that not all infinite sets are countable. Often it turns out to be important to find whether a particular set M is finite or countable, or whether it is infinite but uncountable.

Ex. 12.1. Show that the rationals are countable by putting $x_{jk} = j/k$, obtaining a sequence $\{x_n'\}$ of distinct numbers, and then considering $0, x_1', -x_1', x_2', -x_2', \ldots$

Ex. 12.2. Show that the irrationals are not countable.

13. The real line

The basic result of real variable theory is as follows.

THEOREM 13.1. *A non-empty set* S *of real numbers that is bounded above by a number* b, *so that* x ⩽ b *for all* x *in* S, *has a supremum (least upper bound)* u, *i.e. there is a real number* u *such that* x ⩽ u *for all* x in S; *and if* v < u, *there is an* x *of* S *in* v < x ⩽ u. *Clearly* u ⩽ b. *A non-empty set* S *of real numbers that is bounded below by a number* c, *so that* x ⩾ c *for all* x *in* S, *has an infimum (greatest lower bound)* w, *i.e. there is a real number* w *such that* x ⩾ w *for all* x *in* S; *and if* v > w, *there is an* x *of* S *in* v > x ⩾ w. *Clearly* w ⩾ c.

We either prove this theorem from the particular construction of the real numbers used, or we can take it as an axiom. The latter is the course we adopt here. The second half of the theorem can be deduced from the first half on replacing x by $-x$.

limits here.

THEOREM 13.2. *Let* g *be a bounded above and monotone increasing function of* u, *i.e.* g(u) ⩽ g(v) ⩽ b *for some* b *and all* u < v *in the range considered, which can be the positive integers (when* g(u) *forms a real sequence) or the positive real numbers. Then there exists*

$$\lim_{u\to\infty} g(u) \leqslant b$$

Let S_1 be the set of $g(u)$ for all u in the given range. Then S_1 is non-empty and is bounded above by b. Hence by Theorem 13.1 it has a supremum, say, g_1, such that $g(u) \leqslant g_1$ for all u in the range, and if $g_2 < g_1$ there is a v of the range such that

$$g_2 < g(v) \leqslant g_1$$

Then for all w in $w \geqslant v$ we have by monotonicity,

$$g_2 < g(v) \leqslant g(w) \leqslant g_1$$

and the theorem follows.

There are several variants of this theorem. For example, g can be bounded below and be monotone decreasing,

$$g(u) \geqslant g(v) \geqslant b \ (u < v)$$

and we obtain the existence of the limit. Further, for suitable ranges of u the limit can be

$$\lim_{u\to-\infty}, \quad \lim_{u\to a-}, \quad \lim_{u\to a+}$$

The last two limits are taken as $u \to a$, with $u < a, u > a$, respectively; and if applied to g, the respective limits are denoted by $g(a-)$, $g(a+)$.

Two real numbers $u < v$ define four intervals, namely, $u \leqslant x \leqslant v$, written $[u, v]$ and called *closed*; $u < x < v$, written (u, v) and called *open*; $u < x \leqslant v$, written $(u, v]$ and called *open on the left*; and $u \leqslant x < v$, written $[u, v)$ and called *open on the right*. An x in $u < x < v$ is said to be *inside* all four intervals. An *elementary set* is a union of a finite number of disjoint closed intervals. Two intervals or elementary sets are *overlapping* (*non-overlapping*) if they have (do not have) an open interval in common. Thus two intervals can Sometimes be non-overlapping but not disjoint, having a common end-point.

A division $\mathscr{D}$ of an elementary set E is a finite number of disjoint intervals $[u, v)$ contained in E, such that their *closures* $[u, v]$ have union E. We sometimes use the notation

$$\mathscr{D}\{u, v\}$$

to point out the general interval $[u, v)$ of $\mathscr{D}$. This apparently arbitrary choice of intervals open on the right is made to avoid repeated remarks about common end-points v of intervals defined from $u < v, v < w$. For

$$[u, v) \cup [v, w) = [u, w),$$

a union of two disjoint sets; whereas, for example, the first set of

$$[u,v) \cup (v,w), \quad [u,v] \cup [v,w)$$

has v missing, while the second is a union of non-disjoint sets. We could alternatively have used intervals open on the left, in every division in the book. Further, we use finite unions E of *closed* intervals because we often need to consider the set S_2 of end-points of intervals of all divisions of E, and $S_2 = E$.

We now have the covering theorem of E. Borel (1871–1956).

THEOREM 13.3. *Let $\mathscr{F}$ be a family of open intervals such that each point of a closed interval $[a, b]$ is inside an interval of $\mathscr{F}$. Then a finite family $\mathscr{F}_1$ of the intervals of $\mathscr{F}$ will cover $[a, b]$, i.e. each point of $[a, b]$ will be inside one or more of the intervals of $\mathscr{F}_1$.*

Let S_3 be the set of all points x in $[a, b]$ such that $[a, x]$ (if $x > a$) or a (if $x = a$) can be covered by a finite number of intervals of $\mathscr{F}$. The point a is in an interval of $\mathscr{F}$, so that S_3 contains a and is not empty. It is bounded above by b, and so by Theorem 13.1 has a supremum u in $a \leqslant u \leqslant b$. By hypothesis u is in an interval of $\mathscr{F}$, say $(u-2v, u+2w)$, for some $v > 0$, and $w > 0$, and $u-v$ is in S_3 or is less than a. Adding $(u-2v, u+2w)$ to the finite collection covering $[a, u-v]$ (when $u-v > a$), or using $(u-2v, u+2w)$ alone (when $u-v \leqslant a$) we see that the smaller of b and $u+w$ is in S_3. This contradicts the definition of u, unless $u = b$ and the theorem is proved.

Ex. 13.1. Prove (2.5) from $D(H-H_1) = 0$, by using Theorem 13.3.

The following sections 14 and 15 are not needed in the bulk of the theory. Their only use is here but they are vital sections in Lebesgue measure theory in the explanation of the Denjoy extension of Chapter 7, and in Theorem 49.1 (p. 121).

*14. Open sets

A set G of real numbers is *open* if each point of G is in an open interval wholly contained in G.

III D.

THEOREM 14.1. *An open set is the union of a finite or countable number of disjoint open intervals.*

For fixed x in G let S be the set of points u such that (x, u) is in G. By definition S is not empty. If S is unbounded above then, given any $u > x$, there is a point w of S greater than u, so that u is in S, and $(x, +\infty)$ is in G. If S is bounded above then it has a supremum v. If $x < u < v$ there is a w of S in $u < w \leqslant v$, so that u is in S, and (x, v) is in G. If v is in G, there is an interval (u, w) in G with $u < v < w$, by definition, so that w is in S, and yet $w > v$. Hence v is not in G. Similarly, for some $t < x$, (t, x) is in G but t is not in G. An interval (u, v) in G, with u, v not in G, can be called a *component* interval of G. Thus we have proved that *each point of G lies in a component interval of G.* Two distinct component intervals of G cannot overlap, or else one of the end-points lies in the union of the intervals and so in G, contrary to hypothesis.

$t(x) = \frac{x}{1-x^2}$

By the transformation $y = t(x)$ of Section 47 (p. 115) a real number y transforms into an x between 1 and -1, an open interval into a finite open interval, and disjoint intervals into disjoint intervals. Thus we can suppose for simplicity that G lies in $(-1, +1)$. Let G have m component intervals of length greater than $1/n$. Their total length cannot exceed 2, so that we have $m/n < 2$, $m < 2n$. Thus we can put in order from left to right the intervals with length greater than $\frac{1}{2}$, and, after them, put in order from left to right, the intervals with length between $\frac{1}{2}$ and $\frac{1}{3}$, and so on. At each stage we have only a finite number of intervals, so that we have proved the theorem.

III D. THEOREM 14.2. *A union and a finite intersection of open sets are open.*

For suffices α that need not be countable, let G_α be open, and put

$$U = \bigcup G_\alpha$$

If x is in $\tilde{U}$, then there is at least one α for which x is in G_α, so that x is in an open interval wholly contained in G_α, and so in U. Thus U is open.

If the α are restricted to values $j = 1, 2, \ldots, n$, with

$$W = \bigcap_{j=1}^{n} G_j,$$

and if x is in W, then x is in G_j for $j = 1, 2, \ldots, n$. Hence x is in an open interval J_j wholly contained in G_j. The intervals $J_1, \ldots, J_n$ have an open interval J in common that includes x. As J_j is wholly contained in G_j, so is J, so that J is wholly contained in W, and W is open.

Ex. 14.1. Find an example to show that the second part of Theorem 14.2 is false for even a countable infinity of open sets G_j.

(The open intervals $(-1/j, +1/j)$ have intersection containing 0 only.)

*15. Limit-points and closed sets

III0 A point x is a *limit-point* of a set X of points on the real axis if, for each $\epsilon > 0$, there are points of X other than x in the interval $(x-\epsilon, x+\epsilon)$. *An isolated point of* X is a point of X that is not a limit point of X. The set X' of all limit-points of X is called the *derived set of* X, and the union of X and X' is $\overline{X}$, *the closure of* X. To show that this definition of derived set agrees with that given in Section 33 (p. 71) we have only to observe that if x is in X' there is a point x_1, not x, in $(x-1, x+1) \cap X$, and then a point x_2, not x, in $(x-|x-x_1|, x+|x-x_1|) \cap X$, and so on, so that the point x lies in the derived set of X according to the definition in Section 33 (p. 71). The converse is clear.

A set F that contains all its limit-points ($F' \subseteq F, \overline{F} = F$) is closed. The letter F is from *fermé*, French for closed. A set X is *dense-in-itself* if every point of X is a limit-point of $X (X \subseteq X')$. A set P that is closed and dense-in-itself ($P = P'$) is called *perfect*, and hence the P.

III0 THEOREM 15.1. *The complement of an open set is closed, that of a closed set is open.*

Let G be open, and let x be a limit-point of CG. If x is in G there is an $\epsilon > 0$ such that $(x-\epsilon, x+\epsilon)$ is in G, and contains no point of CG. Thus x cannot be in G, and so is in CG. Thus CG contains all its limit-points, and so is closed.

Let F be closed, and let x be in CF. As F' is contained in F, x cannot be in F', so that for some $\epsilon > 0$ there is an interval $(x-\epsilon, x+\epsilon)$ that contains no point of F, and so is wholly within CF. Thus CF is open.

III0. THEOREM 15.2. *An intersection and a finite union of closed sets are closed.*

These follow from Theorems 15.1 and 14.2 on taking complements. For

$$C\bigcap F_\alpha = \bigcup CF_\alpha, \quad C\bigcup_{j=1}^{n} F_j = \bigcap_{j=1}^{n} CF_j$$

Ex. 15.1. Give an example proving that the second part of Theorem 15.2 is false for a countable infinity of closed sets F_j.

Ex. 15.2. Prove Theorem 15.2 without taking complements. (Let x be in X'_1 where $X_1 = \bigcap F_\alpha$.

Then given $\epsilon > 0$, there is a y in X_1 such that

$$0 < |y-x| < \epsilon$$

Then for each α, y is in F_α for each $\epsilon > 0$, and x is in $F_\alpha' \subseteq F_\alpha$. Thus x is in X_1. If x is in X'_2 where

$$X_2 = \bigcup_{j=1}^{n} F_j$$

there is a sequence $\{x_k\}$ in X_2 that tends to x. Then for some $j = 1, \ldots, n$, an infinite subsequence of $\{x_k\}$ lies in F_j, so that x lies in $F'_j \subseteq F_j$, and $X'_2 \subseteq X_2$.)

16. Left-complete and right-complete families

A family $\mathscr{L}$ of intervals $[t, x)$ *is left-complete in* $[v, w]$, if to each x in $v < x \leqslant w$ there is a $\delta_1(x) > 0$ such that $[t, x)$ is in $\mathscr{L}$ for all t in $x-\delta_1(x) \leqslant t < x$. Then $[x-\delta_1(x), x)$ will be called the *defining interval at* x *of* $\mathscr{L}$, and x will be called the *associated point* of each interval $[t, x)$ in $\mathscr{L}$. The family $\mathscr{L}$ *is left-complete in an elementary set* E *if* $\mathscr{L}$ is left-complete in each closed interval contained in E. A family of intervals $[x, u)$ *is right-complete in* $[v, w]$ if to each x in $v \leqslant x < w$ there is a $\delta_2(x) > 0$ such that $[x, u)$ is in $\mathscr{R}$ for all u in $x < u \leqslant x+\delta_2(x)$. The further definitions follow as for $\mathscr{L}$.

Using these kinds of families we are able to define an integration of Riemann type, the crucial theorem being as follows.

THEOREM 16.1.† *If $\mathscr{L}$ is left-complete, and $\mathscr{R}$ right-complete, in an elementary set* E, *there is a division of* E *formed from the intervals of $\mathscr{L}$ and $\mathscr{R}$.*

The proof for E follows from the proof for each closed interval of E. It is thus no restriction to assume that $E = [a, b]$. Then each point x of $[a+\delta_2(a), b-\delta_1(b)]$ is in an interval

$$I(x) = (x-\delta_1(x), x+\delta_2(x))$$

formed from the interior of an interval of $\mathscr{L}$, together with an interval of $\mathscr{R}$. By Theorem 13.3 a finite number of intervals, say, $I(x_1), \ldots, I(x_n)$, together cover the interval $[a+\delta_2(a), b-\delta_1(b)]$. Then $I(x_j)$ must overlap with $I(x_{j+1})$ for $j = 0, 1, 2, \ldots, n$, if we put

$$I(x_0) = [a, a+\delta_2(a)), I(x_{n+1}) = (b-\delta_1(b), b]$$

Putting $a = x_0$, $b = x_{n+1}$, if

$$x_j < t_j < x_{j+1}, t_j \epsilon I(x_j), t_j \epsilon I(x_{j+1}) (j = 0, 1, 2, \ldots, n) \qquad (16.1)$$

then

$$[x_j, t_j) \epsilon \mathscr{R}, [t_j, x_{j+1}) \epsilon \mathscr{L} (j = 0, 1, 2, \ldots, n)$$

These intervals form the required division.

Note that it may sometimes be possible to take $t_j = x_{j+1}$, in which case we can omit the interval $[t_j, x_{j+1})$; and similarly if it is possible to take‡ $t_j = x_j$.

From this point onwards we denote the pair $\{\mathscr{L}, \mathscr{R}\}$ by **A**; or, if we use a suffix σ, we denote the pair $\{\mathscr{L}_\sigma, \mathscr{R}_\sigma\}$ by $\mathbf{A}_\sigma$. Also, if we say that $\mathbf{A}_\sigma$ *is complete in E*, a one-dimensional set, we shall mean that

$$\mathbf{A}_\sigma = \{\mathscr{L}_\sigma, \mathscr{R}_\sigma\}$$

where $\mathscr{L}_\sigma$ is left-complete, and $\mathscr{R}_\sigma$ right-complete, in the elementary set E. This will simplify many statements, and will be useful for those who would like to connect this book with ~~ref. 13~~ Henstock (1961a). Two-dimensional sets are discussed in Chapter 6.

---

† This was discovered independently of the paper by N. Lusin (1911-12). See Section 50 (p. 124).

‡ Theorem 16.1 shows that the pair $\{\mathscr{L}, \mathscr{R}\}$ *is complete in E*, in the sense of ~~the paper given in~~ Henstock (1961a). This book can be read without a knowledge of that paper.

An extension of theorem 16.1 is as follows.

THEOREM 16.2. *Let* M *be a set of points everywhere dense*† *in* [a, b]. *Then we can restrict the points* $t_j (j = 0, 1, 2, \ldots, n)$ *of Theorem* 16.1 *to lie in* M. *Similarly for an elementary set* E.

To conditions (16.1) we can add $t_j \in M$.

The only difference between Riemann, and the Riemann-complete integration to be defined in the next section, lies in the use of complete sets **A**, and these are crucial to our theory. We therefore consider a few simple properties and examples.

Ex. 16.1. If $\mathscr{L}$ is left-complete in an elementary set E, and if E_1 is an elementary set contained in E, prove that $\mathscr{L}$ is left-complete in E_1. Similarly for $\mathscr{R}$.

Ex. 16.2. If $\mathscr{L}_1$ contains $\mathscr{L}$, which is left-complete in E, prove that $\mathscr{L}_1$ is left-complete in E. Similarly for $\mathscr{R}_1$ containing $\mathscr{R}$.

Ex. 16.3. We can take $\delta_1(x) = \delta_2(x) > 0$.

Ex. 16.4. If **A** is complete in E, and if E_1 is an elementary set contained in E, prove that there is a division of E_1 formed from the intervals of $\mathscr{L}$, $\mathscr{R}$.

Ex. 16.5. Let $\mathscr{F}$ be the family of all intervals $[u, v)$ in [0, 1] with $\frac{3}{2}u < v \leqslant 2u$. Then $\mathscr{F}$ contains no interval with $u = 0$, and contains arbitrarily small intervals. But if $\mathbf{A}_2$ is defined by

$$0 < \delta_1(x) \leqslant \tfrac{1}{3}x, 0 < \delta_2(x) \leqslant \tfrac{1}{2}x (x > 0), \delta_2(0) > 0$$

show that $\mathbf{A}_2$ does not use any interval of $\mathscr{F}$.

Ex. 16.6. Generally, let $\mathscr{F}_n$ be the family of all $[u, v)$ in [0, 1] with

$$u(n+1)/n < v \leqslant un/(n-1) \ (n \geqslant 2)$$

If $\mathbf{A}_{3n}$ is defined by

$$0 < \delta_1(x) \leqslant x/(n+1), \ 0 < \delta_2(x) \leqslant x/n (x > 0), \ \delta_2(0) > 0$$

show that $\mathbf{A}_{3n}$ does not use any interval of $\mathscr{F}_n$.

Ex. 16.7. Show that if the $\mathscr{F}_n$ are as in Ex. 16.6, every complete set in [0, 1] must use intervals of

$$\bigcup_{n=2}^{\infty} \mathscr{F}_n$$

Ex. 16.8. If $\mathbf{A}_4$, $\mathbf{A}_5$ are both complete in E, then so is

---

† That is, to each x in $[a, b]$ there corresponds a sequence $\{m_j\}$ of points of M that tends to x.

$\{\mathscr{L}_4 \cap \mathscr{L}_5, \mathscr{R}_4 \cap \mathscr{R}_5\}$. We denote the latter complete set by $\mathbf{A}_4.\mathbf{A}_5$, *the product of* $\mathbf{A}_4$ *and* $\mathbf{A}_5$. Then we can write $\mathbf{A}_4 = \mathbf{A}_4.\mathbf{A}_5$ when $\mathscr{L}_4 \subseteq \mathscr{L}_5, \mathscr{R}_4 \subseteq \mathscr{R}_5$.

Ex. 16.9. If $u \neq v$ then $\delta_j(u), \delta_k(v)(j, k = 1, 2)$ can be varied independently. Hence show that if to each x in an elementary set E there corresponds a complete set $\mathbf{A}_6(x)$ in E, we can construct from the $\mathbf{A}_6(x)$ a single $\mathbf{A}_7$ complete in E, with δ_j-functions at x equal to the corresponding δ_j-functions at x of $\mathbf{A}_6(x)$. We can even take *two* $\mathbf{A}_6(x)$ for each x, and use the left-complete family from one, and the right-complete family from the other.

17. The integration of point functions

If $\mathbf{A}$ is complete in E, let Δ be the family of all divisions of E formed from intervals of $\mathscr{L}$ and $\mathscr{R}$, together with a specification attached to each division that states the family from which each interval is taken. Thus if exactly m intervals of a division come from both $\mathscr{L}$ and $\mathscr{R}$ we can obtain 2^m-1 more specifications solely by altering the family from which one or more of the intervals come. By Theorem 16.1, Δ is not empty.

If f is a finite real or complex-valued point function in an elementary set E, and if

$$\mathscr{D}\{u, v\} \in \Delta, \textit{ let } (\mathscr{D}) \sum f(x)(v-u)$$

denote the sum over the division $\mathscr{D}$, where x is the associated point of $[u, v)$, according to the given specification. We denote by $S(f, x; \mathbf{A}; E)$ the family of all such sums, for all $\mathscr{D}$ in Δ. Then the real or complex number I is the Riemann integral using complete sets, or the *Riemann-complete integral, of f in E*, if the sets $S(f, x; \mathbf{A}; E)$ shrink up to I as we use finer and finer $\mathbf{A}$; more exactly, if, given $\epsilon > 0$, then $\mathbf{A}$ can be so chosen that $S(f, x; \mathbf{A}; E)$ lies in the closed circle on the complex plane with centre I and radius ϵ. If the sums are real, then I must be real, and the sets must clearly lie in $[I-\epsilon, I+\epsilon]$. We write I compactly as

$$RC(f, x; E)$$

or, in a more usual form,

$$(RC) \int_E f \, \mathrm{d}x$$

If X is a set on the real axis we denote by $ch(X) = ch(X; x)$

its *characteristic function*, the function that is equal to 1 in X, and to 0 elsewhere. By using this we can *integrate in X, relative to E*, by replacing f by $f.ch(X)$. In this way we can integrate in arbitrary bounded sets. For unbounded sets we use the theory of Chapter 7.

More generally, we can replace x in the definition of Riemann-complete integration by another point function g of x, obtaining the Riemann-complete integral

$$RC(f, g; E) = (RC) \int_E f \, \mathrm{d}g$$

by using sums

$$(\mathscr{D}) \sum f(x)\{g(v)-g(u)\}$$

Ex. 17.1. Show that two different complex numbers cannot have the property of I, for the same f, x, E, so that, if it exists, $RC(f, x; E)$ is uniquely defined.

Ex. 17.2. If H is the indefinite Newton integral of a finite function f in $[a, b]$, show by using Section 4 that the Riemann-complete integral of f in $[a, b]$ exists and is equal to

$$H(b)-H(a)$$

Ex. 17.3. By using (3.1) to (3.4) and Section 4, and taking

$$0 < \delta_1(x) < \delta, \quad 0 < \delta_2(x) < \delta$$

show that if the Riemann integral of the bounded function f exists then so does $RC(f, x; [a, b])$, and they are equal.

Ex. 17.4. If the Riemann–Stieltjes integral of f with respect to g exists in $[a, b]$, prove that $RC(f, g; [a, b])$ exists and is equal to it.

* *Ex. 17.5.* If the Pollard–Getchell integral of f with respect to g exists in $[a, b]$, prove that the corresponding Riemann-complete integral exists and is equal to it.

18. Functions of intervals

The intervals in our divisions are all open on the right. A function h of such intervals, or an *interval function*, is defined in an elementary set E if to each interval in E and open on the right there is assigned a number. We write $h(u, v)$ in place of $h([u, v))$. In Section 17 we first used the interval functions

$$f(v)(v-u), \quad f(u)(v-u)$$

and then the interval functions

$$f(v)\{g(v)-g(u)\}, f(u)\{g(v)-g(u)\}$$

A most fruitful generalization is to use two interval functions, one to be associated with intervals from an $\mathscr{L}$, and one with intervals from an $\mathscr{R}$. The definitions are as follows.

If h_l, h_r are two interval functions we denote the pair by $\{h_l, h_r\}$ or by $\mathbf{h}$. If $\mathbf{A}$ is complete in E, and if Δ is obtained from $\mathbf{A}$, and if

$$\mathscr{D}\{u, v\} \in \Delta, \text{ let } (\mathscr{D}) \sum h_s(u, v) = (\mathscr{D}) \sum h_s$$

denote the sum of h_s over the division $\mathscr{D}$, where $s = l$ if $[u, v)$ is counted to be from $\mathscr{L}$, and $s = r$ if $[u, v)$ is from $\mathscr{R}$. We denote by

$$S(\{h_l, h_r\}; \mathbf{A}; E) \text{ or } S(\mathbf{h}; \mathbf{A}; E)$$

the collection of all such sums, for all $\mathscr{D}$ in Δ. Then the real or complex number I is the *Riemann-complete integral of* $\mathbf{h}$ *in* E if, given $\epsilon > 0$, we can choose $\mathbf{A}$ so that $S(\mathbf{h}; \mathbf{A}; E)$ lies in the closed circle with centre I and radius ϵ. The notations for I are

$$RC(\mathbf{h}; E) = RC(\{h_l, h_r\}; E) = (RC)\int_E \{h_l, h_r\} = (RC)\int_E \mathbf{h}$$

For simplicity, when $h_l = h_r = h$, we replace $\{h,h\}$ by h. Also, if f is a point function the pair of interval functions

$$f(v)h_l(u, v), f(u)h_r(u, v)$$

are written more simply as $\{fh_l, fh_r\}$, and the more usual notations for the integral are

$$(RC)\int_E f\,d\{h_l, h_r\} = (RC)\int_E f\,d\mathbf{h}$$

There are no clashes of notation. For in $S(f, g; \mathbf{A}; E)$, $RC(f, g; E)$, the f and g are point functions. In $S(\mathbf{h}; \mathbf{A}; E)$, $S(\{h_l, h_r\}; \mathbf{A}; E)$, $RC(\{h_l, h_r\}; E)$, $RC(\mathbf{h}; E)$, the h_l, h_r are interval functions. In $S(h; \mathbf{A}; E)$, $RC(h; E)$ the h is an interval function. Finally, in $S(\{fh_l, fh_r\}; \mathbf{A}; E)$, $RC(\{fh_l, fh_r\}; E)$, the f is a point function and the h_l, h_r are interval functions.

Ex. 18.1. If the Burkill integral of h exists in E, then $RC(h; E)$ exists and is equal to it. (Follow Ex. 17.3 and Ex. 17.4.)

Notation $\int_E f(P)dP$ due to Leb. "Sur l'intégration des fonctions discontinues" (Ann. Éc. Norm., 1910)

19. Distributive functional

A *functional* of the pair $\mathbf{h}$ of interval functions is defined in a family $\mathscr{B}$ of pairs, if to each pair in $\mathscr{B}$ there is assigned a real or complex number. The functional p *of* $\mathbf{h}$ *is distributive* if for arbitrary real or complex constants α, β, and arbitrary pairs $\mathbf{h}, \mathbf{h}^*$ from $\mathscr{B}$, we have

$$p(\{\alpha h_l + \beta h_l^*, \alpha h_r + \beta h_r^*\}) = \alpha p(\{h_l, h_r\}) + \beta p(\{h_l^*, h_r^*\})$$

Ⓝ
III Th. 4.4

THEOREM 19.1. *Let $\mathscr{B}$ be the family of all pairs of interval functions that have a Riemann-complete integral in an elementary set E. If α, β are constants, and if $\mathbf{h}, \mathbf{h}^*$ are in $\mathscr{B}$, then so is $\{\alpha h_l + \beta h_l^*, \alpha h_r + \beta h_r^*\}$, and the Riemann-complete integral is a distributive functional in $\mathscr{B}$.*

✓ To each $\epsilon > 0$ there correspond $\mathbf{A}, \mathbf{A}_1$, complete in E, such that if $\mathscr{D}$, $\mathscr{D}_1$ are from $\mathbf{A}, \mathbf{A}_1$, respectively, and

$$I = RC(\mathbf{h}; E), I^* = RC(\mathbf{h}^*; E)$$

then

Th. 23.2

$$|I - (\mathscr{D})\sum h_s| \leqslant \epsilon, \quad |I^* - (\mathscr{D}_1)\sum h_s^*| \leqslant \epsilon$$

Further, by Ex. 16.8, $\mathbf{A}_2 = \mathbf{A}.\mathbf{A}_1$ is complete in E. If $\mathscr{D} = \mathscr{D}_1$ is taken from $\mathbf{A}_2$,

$$|\alpha I + \beta I^* - (\mathscr{D})\sum \{\alpha h_s + \beta h_s^*\}| \leqslant \epsilon(|\alpha| + |\beta|)$$

Hence the result.

III Th. 4.6
Ⓝ
Th. 23.2

THEOREM 19.2. *If $h_s, h_s^*(s = l, r)$, I, I^* are all real and $I + iI^*$ is the Riemann-complete integral of $\{h_l + ih_l^*, h_r + ih_r^*\}$ in E, then I, I^* are the respective Riemann-complete integrals of $\mathbf{h}, \mathbf{h}^*$ in E.* ✓

✓ Given $\epsilon > 0$, there is a set $\mathbf{A}_3$ complete in E such that every division $\mathscr{D}_2$ from $\mathbf{A}_3$ satisfies

$$\epsilon \geqslant |I + iI^* - (\mathscr{D}_2)\sum \{h_s + ih_s^*\}|$$
$$= |I - (\mathscr{D}_2)\sum h_s + i\{I^* - (\mathscr{D}_2)\sum h_s^*\}| \geqslant |I - (\mathscr{D}_2)\sum h_s|$$

Thus I is the Riemann-complete integral of $\mathbf{h}$ in E, and similarly for I^*. This is a partial converse of Theorem 19.1.

20. Order relation

After addition, and multiplication by constants, we next examine the relation $\leqslant$ for real interval functions.

McGrotty on lattices.
Mine on Φ-function later

So ∫ → Taylor better than Taylor → ... u.c., so diff
a reg in more intuitive than a power series!
$(x_1, x_2, \ldots)$

$x_n = \frac{1}{n}$

P 412 Line 23	$= \overline{\Sigma}_n$	$= \overline{\overline{\Sigma}}_n$

Hilbert matrix if $\Sigma\|a_{mn}\|^2$ is C.

P.437 line 13 (a.)	$\lim_{z \to e^{i\theta}}$	$\lim_{z \to e^{i\theta}}$
.. (diagram) $e^{i\phi}$	$e^{i\theta}$	
P.439 line 7	$\|s_n(e^{i\theta})$	$\|s_n(e^{i\theta})\|$
P.480 .. 12	$s^{(r)}_{,n}(z)$	$s^{(r)}_{\alpha,n}(z)$
P100 .. 28	$n > a^2$	$n > a^{a^2}$
P101 .. 18	$< \frac{2}{\|1-z\|}$	$\leq \frac{2}{\|1-z\|}$
P110 .. 11 (f: b:)	i sin ix	−i sin ix
P.204. .. 6	$-M[f(b)-f(a)]$	$-M[g(b)-g(a)]$
.. 8	$M[f(b)-f(a)]$	$M[g(b)-g(a)]$
.. 24	are not greater	are not less ~~greater~~. ~~[illegible]~~

From **DEIGHTON, BELL & CO. Ltd.**
13 TRINITY STREET, CAMBRIDGE.

In reply to your order dated..
for :

Youngs:- Theory of Integration

P 345	line 18	$\sum_{k=1} g_k(a)$	$\sum_{k=1}^{\infty} g_k(a) c_k$
P 396	" 12	is	there is
"	" 2 (f.b.)	$i > k$	$i \geq 2$
P 404	" 21	$f_k(a) > 0$	$f_k(a) \geq 0$

We regret that we are unable to supply for the reason indicated below (Marked with a X.)

.......X.......OUT OF PRINT.
...................TEMPORARILY OUT OF PRINT.
...................OUT OF STOCK.
...................NOT TO BE RE-ISSUED TILL AFTER THE WAR.
...................REPRINTING. *No date.*
...................BINDING. *No date.*
...................NEW EDITION IN PREPARATION. *No date.*
...................AMERICAN IMPORTATION. *Awaiting supplies from U.S.A. No date.*
...................NOT YET PUBLISHED.
...................SHALL WE TRY TO OBTAIN A SECOND-HAND COPY?
...................WE WILL SUPPLY AS SOON AS AVAILABLE.

Tel. 3939.

P.17 line 5	ϵ_n	η_n
P 20 ·· (from bottom)	I_{n+1}	I_{n-1}
P 22 ·· 3	point of S	point of S other than (a, a')
·· 2 (f: b:)	$\omega+1$	ω
	$S^{\omega+1}$	S^{ω}
P 11 ·· 3	$1+\frac{n-1}{n}$	$1+\frac{m-1}{m}$
P 29 ·· 24	$\theta(\delta_n, f)$	$\theta_c(\delta_n, f)$
P 40 ·· 26	$j_i \leq$	$j_i >$
P 74 ·· 5 (f.b.)	so that	(Delete altogether.)
	both	Both
P 168 ·· 1 (f.b.)	$D\|/\|C\|$	$\|D\|/\|C\|$
P 195 ·· 5 (··)	$\epsilon < 0$	
P 262 ·· 6	$F(z)\|$ $z\|$	$\|F(z)\|$ $\|z\|$
P.352 ·· 23	$\|\frac{z}{\sigma}\|$	$\|\frac{z}{\sigma}\|$
P.452 ·· 19	$\left(\frac{\sin nt}{\sin t}\right)^{\bullet}$	$\left(\frac{\sin nt}{\sin t}\right)^2$
P 262 ·· 15	$f(z'/R)$	$F(z'/R)$
·· 21	$\left\|\frac{v}{Ra}\right\|$	$\left\|\frac{v}{Ra_1}\right\|$
P 268 ·· 21	$-\sqrt{\frac{\log f(z)-1}{2\pi i}}$	$\sqrt{\frac{\log f(z)-2\pi i}{2\pi i}}$
P.38 ·· 9	is increasing	is strictly increasing.
·· ·· 9 (f: b:)	$f(y_n) > f(x_n)$	$f(y_n) \geq f(x_n)$
P.42 ·· 6 (··)	were	where.
P.387 ·· 4	[illegible]	[illegible]

Th 4.5(4.5)

Th 2.3.3

Theorem 20.1. *If* $h_l \leqslant h_l^*$ *in* $\mathscr{L}$, $h_r \leqslant h_r^*$ *in* $\mathscr{R}$, *where* A *is complete in* E, *and if* **h**, **h*** *are Riemann-complete integrable in* E, *then*

$$\mathrm{RC}(\mathbf{h}; \mathrm{E}) \leqslant \mathrm{RC}(\mathbf{h}^*; \mathrm{E}). \tag{20.1}$$

If in the same $\mathscr{L}$, $\mathscr{R}$

$$h_s^j \geqslant 0 \ (j = 1, 2, \ldots), \ h_s = \sum_{j=1}^{\infty} h_s^j \ (s = l, r)$$

where the $\{h_l, h_r\}$, $\{h_l^j, h_r^j\}$ *have Riemann-complete integrals in* E, *then*

$$\mathrm{RC}(\{h_l, h_r\}; \mathrm{E}) \geqslant \sum_{j=1}^{\infty} \mathrm{RC}(h_l^j, h_r^j\}; \mathrm{E}) \tag{20.2}$$

(Here, the h_s^j are not *powers* of h_s.)

Let I, I^* be the first and second integrals, respectively, in (20.1). Then given $\epsilon > 0$, there are A_1, A_2, complete in E, such that every division $\mathscr{D}$ from A_1, $\mathscr{D}_1$ from A_2, satisfy

$$|I-(\mathscr{D})\sum h_s| \leqslant \epsilon, \ |I^*-(\mathscr{D}_1)\sum h_s^*| \leqslant \epsilon$$

As by Ex. 16.8, $A_3 = A.A_1.A_2$ is complete in E, we can take $\mathscr{D} = \mathscr{D}_1$ from A_3 and obtain

$$I-\epsilon \leqslant (\mathscr{D})\sum h_s \leqslant (\mathscr{D})\sum h_s^* \leqslant I^*+\epsilon$$

As $\epsilon > 0$ is arbitrary we obtain the first result.

For (20.2) we use (20.1) and Theorem 19.1, obtaining for each n,

$$h_s \geqslant \sum_{j=1}^{n} h_s^j \ (s = l, r)$$

$$RC(\{h_l, h_r\}; E) \geqslant RC(\{\sum_{j=1}^{n} h_l^j, \sum_{j=1}^{n} h_r^j\}; E) = \sum_{j=1}^{n} RC(\{h_l^j, h_r^j\}; E)$$

Then (20.2) follows on letting $n \to \infty$.

Strict inequality can occur in (20.2); see Ex. 20.1. Further, it is necessary to assume that $\{h_l, h_r\}$ is Riemann-complete integrable in E, as this cannot be deduced from the other hypotheses. See Ex. 20.2.

Thus theorems on the interchange of integral and limit need special pairs of interval functions. Results given in Chapters 4 and 5 cover all the cases of Lebesgue theory.

✓ Ex. 2.3.5

Ex. 20.1. Using Ex. 16.6 we put

$$h_l{}^j(u, v) = h_r{}^j(u, v) = \begin{cases} v-u \ ([u, v] \in \mathscr{F}_j, j > 1) \\ 0 \ \textit{otherwise} \end{cases}$$

Then we have

$$h_l(u, v) = h_r(u, v) = \begin{cases} v-u \ (u < v \leqslant 2u) \\ 0 \ \textit{otherwise} \end{cases}$$

$$RC(\{h_l{}^j, h_r{}^j\}; [0, 1]) = 0 (j = 1, 2, \ldots), RC(\{h_l, h_r\}; [0, 1]) = 1$$

✓ Ex. 2.3.6

Ex. 20.2. If in Ex. 20.1 we put

$$h_{l1} = h_{r1} = \sum_{j=1}^{\infty} h_l^{2j}$$

then $\{h_{l1}, h_{r1}\}$ is not Riemann-complete integrable in any interval of $[0, 1]$. Given a set A_4 complete in $[0, 1]$, let S be the set of all x in $[0, 1]$ such that either $x = 1$ or the supremum of finite sums of h_s over divisions from A_4 over $[x, 1]$ is $1-x$. If α is the infimum of S then $0 \leqslant \alpha \leqslant 1$. Also α is in S. If $\alpha > 0$, there is an integer n such that the interval

$$[(2n-1)\alpha/(2n), \alpha)$$

of $\mathscr{F}_{2n}$ is in $\mathscr{L}_4$, and then $(2n-1)\alpha/(2n)$ is in S, contrary to the definition of α. Hence $\alpha = 0$, and the supremum of finite sums of h_s over divisions from A_4 over $[x, y]$ is $y-x$, for $0 \leqslant x < y \leqslant 1$. Similarly, replacing $2n$ by $2n+1$, we prove that the infimum is 0 and the Riemann-complete integral in $[x, y]$ cannot exist for $0 \leqslant x < y \leqslant 1$.

✓ Th 2.3

Ex. 20.3. If on X, $m = \inf f$ and $M = \sup f$, while $f = 0$ on CX, if $h_s \geqslant 0 (s = l, r)$ and if $\{ch(X)h_l, ch(X)h_r\}$ is Riemann-complete integrable in E, then the Riemann-complete integral of $\{fh_l, fh_r\}$ in E, if it exists, lies between m and M times the Riemann-complete integral of $\{ch(X)h_l, ch(X)h_r\}$. (For $mh_s \leqslant fh_s \leqslant Mh_s$ on X.)

Ex. 20.4. If, in Ex. 20.3, f is continuous and $E = [a, b]$, then there is a ξ in $a \leqslant \xi \leqslant b$ such that

$$RC(\{fh_l, fh_r\}; [a, b]) = f(\xi)RC(\{h_l, h_r\}; [a, b]),$$

(Take $X = [a, b]$ also.)

21. Additivity in elementary sets

An interval function h *is finitely additive* in an elementary set E if

$$h(u,v)+h(v,w) = h(u,w)$$
$$(\textit{all } u, v, w \textit{ with } u<v<w \textit{ and } [u,w] \subseteq E) \qquad (21.1)$$

If E_1 is an elementary set in E we can then define $h(E_1)$ to be the sum of $h(I)$ over the finite number of disjoint closed intervals I with union E_1. Then if E_1, E_2 are two non-overlapping elementary sets in E,

$$h(E_1)+h(E_2) = h(E_1 \cup E_2) \qquad (21.2)$$

Alternatively, if h is defined for each elementary set E_1 in E and satisfies (21.2), then *h is finitely additive for elementary sets in E*. If h is finitely additive in $[a, b]$, and if we write

$$f(x) = h(a,x)$$

then for $a<x<y\leqslant b$ we have

$$f(y) = h(a,y) = h(a,x)+h(x,y) = f(x)+h(x,y)$$
$$h(x,y) = f(y)-f(x) \qquad (21.3)$$

We write this difference as $H_f(x, y)$. Result (21.3) is true also for $x = a$ if we define $f(a) = 0$.

The real interval function h *is finitely superadditive in E* if (21.1) is replaced by

$$h(u,v)+h(v,w) \leqslant h(u,w)$$
$$(\textit{all } u, v, w \textit{ with } u<v<w \textit{ and } [u,w] \subseteq E) \qquad (21.4)$$

Cf. $h(E)$ is defined, where E is the union of disjoint intervals $[a_j, b_j]$ $(j=1,\dots,n)$, we

Further, the real *h is finitely subadditive in E* if $-h$ is finitely superadditive in E. also suppose that $h(E) \geqslant \sum h(a_j, b_j)$.

It follows by induction that if $\mathscr{D}\{u, v\}$ is a division of an interval J,

$$(\mathscr{D}) \sum h(u,v) = h(J) \qquad (21.5)$$

when h is finitely additive; if h is finitely superadditive,

$$(\mathscr{D}) \sum h(u,v) \leqslant h(J) \qquad (21.6)$$

while if h is finitely subadditive,

$$(\mathscr{D}) \sum h(u,v) \geqslant h(J) \qquad (21.7)$$

III Th. 5.1

Ⓝ √ THEOREM 21.1. *If* RC($\mathbf{h}$; E) *exists then* RC($\mathbf{h}$; E$_1$) *exists for each elementary set* E$_1 \subseteq$ E, *and is a finitely additive function of such* E$_1$.

Th. 2, 5.2

If $E_1 \subset E$, there is an elementary set E_3, non-overlapping with E_1, such that $E_1 \cup E_3 = E$. If $\mathbf{A}$ is complete in E, if S_1, S_2 are sums over two divisions of E_1, and if S is a sum over a division of E_3, the divisions being obtained from $\mathbf{A}$, then S_1+S and S_2+S are sums over two divisions of E from $\mathbf{A}$, and so lie in $S(\mathbf{h}; \mathbf{A}; E)$. Given $\epsilon>0$, this set lies in the closed circle centre $I = RC(\mathbf{h}; E)$ and radius ϵ, for suitable choice of $\mathbf{A}$. Hence we obtain a condition similar to that for a fundamental sequence, namely,

$$|S_1-S_2| = |(S_1+S)-(S_2+S)| \leqslant 2\epsilon \qquad (21.8)$$

A fundamental sequence of real or complex numbers is convergent. Here we prove that sums such as S_1, S_2 are 'convergent' in the generalized sense that uses complete sets.

If we replace ϵ by 2^{-j} we have to replace $\mathbf{A}$ by $\mathbf{A}_1{}^j$ complete in E, from which we obtain divisions of E_1 and so sums $S_3{}^j$. We ensure that

$$\mathscr{L}_1{}^1 \supseteq \mathscr{L}_1{}^2 \supseteq \mathscr{L}_1{}^3 \supseteq \cdots$$

and similarly for $\mathscr{R}_1{}^j$, by replacing $\mathbf{A}_1{}^2$ by $\mathbf{A}_1{}^1.\mathbf{A}_1{}^2$, $\mathbf{A}_1{}^3$ by $\mathbf{A}_1{}^1.\mathbf{A}_1{}^2.\mathbf{A}_1{}^3$, and so on, if necessary. Then a sum $S_3{}^k$ is a sum $S_3{}^j$ if $j<k$, and from (21.8),

$$|S_3{}^j - S_3{}^k| \leqslant 2^{-j+1} \; (j<k;\, j, k = 1, 2, \ldots) \qquad (21.9)$$

Taking a fixed sequence $\{S_3{}^j\}$, it tends to a limit, say I_1. Then from (21.8; 21.9)

$$|S_3{}^j - I_1| \leqslant 2^{-j+2} \qquad (21.10)$$

where now $S_3{}^j$ is any point in $S(\mathbf{h}; \mathbf{A}_1{}^j; E_1)$. Thus (21.10) shows that the latter set lies in the circle centre I_1 and radius 2^{-j+2}. Hence the Riemann-complete integral of $\mathbf{h}$ exists in E_1 with value I_1. Similarly for E_3, with integral I_3 say. Finally, if from $\mathbf{A}_1{}^j$ we obtain sums $S_3{}^j$ for E_1, $S_4{}^j$ for E_3, then $S_3{}^j+S_4{}^j$ are sums for E; and $S_3{}^j$ is in the circle centre I_1, radius 2^{-j+2}, while $S_4{}^j$ is in the circle centre I_3, radius 2^{-j+2}, and $S_3{}^j + S_4{}^j$ is in the circle centre I and radius 2^{-j}. Hence

$$|I_1+I_2-I| \leqslant 9\cdot 2^{-j}\ (j = 1, 2, \ldots),\ \ I_1+I_2 = I$$

Similarly for each non-overlapping pair of elementary sets contained in E, proving the theorem.

Note that for fixed E and variable X, it follows from Theorem 19.1 that

$$(RC)\int_E ch(X)\,d\mathbf{h}$$

is finitely additive in X whenever the integral is defined.† The given property in Theorem 21.1 corresponds to the finite additivity of the Burkill integral. From this property it is usual to deduce a theorem concerning partial sums. The corresponding theorem here is as follows.

First, let $\mathbf{A}_2$ be complete in E. Then each sum

$$(\mathscr{Q})\sum h_s(u, v)\ (s = l \textit{ if } [u, v)\in\mathscr{L}_2,\ s = r \textit{ otherwise})$$

obtained from $\mathbf{h}$ and a finite collection $\mathscr{Q}$ of disjoint intervals in E from $\mathscr{L}_2 \cup \mathscr{R}_2$, is called a *partial sum in E using* $\mathbf{A}_2$. Let

$$P(\mathbf{h}; \mathbf{A}_2; E)$$

be the set of all partial sums in E using $\mathbf{A}_2$. We call P *the variation set of* $\mathbf{h}$ *in E, using* $\mathbf{A}_2$, the reason for the name is given later.

Secondly, we say that $\mathbf{h}$ *is continuous at x*, if $h_l(t, x) \to 0$ as $t \to x-$, $h_r(x, u) \to 0$ as $u \to x+$; and that $\mathbf{h}$ *is continuous in E*, if the limits are 0 for all x, with t, u restricted to lie in E. Thus if x is not in E, the definition is vacuous; while if x is a boundary point of E, the definition is one-sided.

III Th. 5.3 (wt cty)

NewRC

THEOREM 21.2. *Let there exist* $\mathrm{RC}(\mathbf{h}; \mathrm{E})$, *and put*

$$\mathrm{H}(\mathrm{u}, \mathrm{v}) = \mathrm{RC}(\mathbf{h}; [\mathrm{u}, \mathrm{v}])$$

Given $\epsilon > 0$, there is an $\mathbf{A}$ *complete in* E *such that*

$$\mathrm{P}(\{\mathrm{h}_l-\mathrm{H}, \mathrm{h}_r-\mathrm{H}\}; \mathbf{A}; \mathrm{E}) \tag{21.11}$$

✓ limiting variation set,

---

† Later, we will be able to take E as the whole axis, and then this result will correspond to the finite additivity of the Lebesgue integral in the sets over which it is defined. Complete additivity will come later, as well.

Th. 2.5.4 (2.5.12)

lies in the circle centre 0, *radius* 2ϵ. *Further*, $\{\mathrm{h}_l-\mathrm{H}, \mathrm{h}_r-\mathrm{H}\}$ *is continuous in* E,

$$\lim_{t\to x-}\{\mathrm{H}(t, x)-\mathrm{h}_l(t, x)\} = 0 \textit{ if for some } t < x,\ [t, x) \subseteq \mathrm{E} \tag{21.12}$$

$$\lim_{u\to x+}\{\mathrm{H}(x, u)-\mathrm{h}_r(x, u)\} = 0 \textit{ if for some } u > x,\ [x, u) \subseteq \mathrm{E} \tag{21.13}$$

and if we put $\mathrm{k}_s = \mathrm{h}_s - \mathrm{H}\ (s = l, r)$, *we have*

$$\mathrm{RC}(\{|\mathrm{k}_l|, |\mathrm{k}_r|\}; \mathrm{E}) = 0 \tag{21.14}$$

Let $\mathbf{A}$ be as in Theorem 21.1, and let $\mathscr{Q}$ be a finite family of disjoint intervals in E from $\mathscr{L} \cup \mathscr{R}$. Then E, less the intervals of $\mathscr{Q}$, is a finite union of intervals, together possibly with a finite number of isolated points. Let E_4 be the union of the corresponding closed intervals. By Theorem 21.1, $H(E_4)$ exists, so that there is a division $\mathscr{D}_1$ of E_4 from $\mathbf{A}_3 = \mathbf{A}.\mathbf{A}_3$ complete in E_4, such that

$$|(\mathscr{D}_1)\sum h_s(u, v)-H(E_4)| < \epsilon\ (s = l \textit{ if } [u, v)\in \mathscr{L}_3; \textit{ and otherwise } s = r) \tag{21.15}$$

By (21.15) and Theorem 21.1, as $\mathscr{Q} \cup \mathscr{D}_1$ gives a division of E from $\mathbf{A}$,

$$|(\mathscr{Q} \cup \mathscr{D}_1)\sum h_s(u, v)-H(E)| < \epsilon,$$

$$|(\mathscr{Q})\sum \{h_s(u, v)-H(u, v)\}| < 2\epsilon$$

giving the first result. The continuity part (21.12;21.13) then follows on taking one interval in $\mathscr{Q}$. For (21.14) we first take those intervals of $\mathscr{Q}$ for which the real part k_{s1} of k_s is positive, and then separately those for which k_{s1} is negative, thus obtaining two partial sums. Doing the same for the imaginary part k_{s2} of k_s, we have

$$(\mathscr{Q})\sum |k_{s1}| < 4\epsilon,\ (\mathscr{Q})\sum |k_{s2}| < 4\epsilon,\ (\mathscr{Q})\sum |k_s| < 8\epsilon \tag{21.16}$$

Taking $\mathscr{Q}$ as a division of E we obtain (21.14).

Ex. 21.1. If χ is finitely superadditive, $h_l \leqslant \chi$ in $\mathscr{L}_4$, $h_r \leqslant \chi$ in $\mathscr{R}_4$, where $\mathbf{A}_4$ is complete in E, and $\mathbf{h}$ is Riemann-complete integrable in E, prove that

$$RC(\mathbf{h}; E) \leqslant \chi(E)$$

h(5) *Ex. 21.2.* Let $f(0) = 1$, $f(x) = 0 (x > 0)$, and let

$$K(0, v) = RC(\{fh_l, fh_r\}; [0, v])$$

exist for $v = 1$. Prove that $K(0, v) = K(0, 1)$ for $v > 0$ 0≤v≤1 and hence deduce that

$$\lim_{u\to 0+} h_r(0, u)$$

exists and is equal to $K(0, 1)$.

22. Two integrability results Th. 2.7.2

Th 4.7

Theorem† 22.1. *Let* h *be real and finitely subadditive. If there is an* A *complete in* E *such that* S(h; A; E) *is bounded above, with supremum* s, *then* h *is Riemann-complete integrable in* E *with integral equal to* s. *However, if* S(h; A; E) *is unbounded above, for all* A, *then for each integer* n *there is an* A_{1n} *complete in* E *such that each point of* S(h; A_{1n}; E) *is greater than* n. *If* h *is real and finitely superadditive, and if* A_2 *is complete in* E *such that* S(h; A_2; E) *is bounded below, with infimum* m, *then* h *is Riemann-complete integrable with integral equal to* m. *However, if* S(h; A_2; E) *is unbounded below, for all* A_2, *then for each integer* n *there is an* A_{3n} *complete in* E *such that each point of* S(h; A_{3n}; E) *is less than* $-$n.

See M

✓

The second two results follow from the first two on putting $-h$ for h. Let $\mathscr{D}$ be a division of E from A, and let $\mathscr{L}_4$, $\mathscr{R}_4$ be the families resulting from $\mathscr{L}$, $\mathscr{R}$, respectively, on removing all intervals that do not lie entirely within some interval of $\mathscr{D}$. Then A_4 is complete in E, and if $\mathscr{D}_1$ is a division of E from A_4, then by finite subadditivity we have

$$(\mathscr{D}_1)\sum h \geqslant (\mathscr{D})\sum h \tag{22.1}$$

In the first case we take the second sum in (22.1) to be greater than $s-\epsilon$, for given $\epsilon > 0$, and then every division $\mathscr{D}_1$ from A_4 gives a sum between $s-\epsilon$ and s. Hence the first result. In the second case we take the second sum in (22.1) to be greater than n.

mplete sets
new RC

Theorem 22.2. *Let* g, g_1 *be two point functions that are equal at points of a set* M *everywhere dense*‡ *in* [a, b], *and let* RC(f, g; [a, b]) *and* RC(f, g_1; [a, b]) *exist. Then* Th 7.2.3

III Ex
4.

$$RC(f, g; [a, b]) = RC(f, g_1; [a, b]) + \\ +f(b)\{g(b)-g_1(b)\}-f(a)\{g(a)-g_1(a)\} \tag{22.2}$$

✓

---

† Theorem 22.1 is like a theorem of Burkill for continuous finitely subadditive interval functions.

‡ That is, to each x in $a \leqslant x \leqslant b$ there corresponds a sequence of points of M tending to x.

This shows that if $g = g_1$ at points of M, and if the first integral in (22.2) exists, then either the second does not exist, or else it differs from the first by a known simple function. Theorem 33.3 (p. 71) will complete the result.

By distributivity (Theorem 19.1) we can assume that $g_1 = 0$, and then we apply Theorem 16.2. Here, M is the set of zeros of g. Given A_5 complete in $[a, b]$, there is a division $\mathscr{D}_2$ of $[a, b]$ that uses the points of M for points t_j and has a and b as associated points at the ends, and

$$f(x_{j+1})\{g(x_{j+1})-g(t_j)\}+f(x_{j+1})\{g(t_{j+1})-g(x_{j+1})\}$$
$$= f(x_{j+1})\{g(t_{j+1})-g(t_j)\} = 0,$$
$$(\mathscr{D}_2)\sum f\{g(v)-g(u)\} = f(b)g(b)-f(a)g(a)$$

giving (22.2).

*23. Stieltjes transformation

This section is used in Theorem 53.1 (p. 130).

New RC

Ne
Rc

Theorem 23.1. *Let* φ *be a finite point function that is monotone increasing in* x ⩾ 0, *with* φ(0) = 0, *and put* $\psi(y)$ = inf x *for all* x ⩾ 0 *such that* φ(x) ⩾ y. *Then we say that* ψ *is the inverse function of* φ. *It has the properties*:

$$\psi(0) = 0, \psi \tag{23.1}$$

is finite and monotone increasing, and

$$\psi(y-) = \psi(y) \quad (y > 0) \tag{23.2}$$

Further let

$$h_s(u, v) = 0 \ (s = 1, r) \ \textit{if} \ \varphi(u+) = \varphi(v) \tag{23.3}$$

$$h_s^*(y, z) = 0 \ (s = 1, r) \ \textit{if} \ \psi(y+) = \psi(z) \tag{23.4}$$

$$h_s^*(y, z) = h_s(\psi(y+), \psi(z)) \ (s = 1, r) \ \textit{otherwise.} \tag{23.5}$$

Then

$$RC(h; [\psi(a), \psi(b)]) = RC(h^*; [a, b]) \tag{23.6}$$

in the sense that if the first integral exists so does the second, and they are equal.

If we add the condition

$$\varphi(x-) = \varphi(x) \quad (x > 0) \tag{23.7}$$

then the definition is symmetrical. The inverse function of ψ is φ.

Results (23.1) are obvious. To *prove* (23.2) let $\psi(y-) = v$. Then for all $z < y$ we have $\psi(z) \leqslant v$, so that $\inf x \leqslant v$ for all $x \geqslant 0$ such that $\varphi(x) \geqslant z$. Hence by monotonicity, $\varphi(v+) \geqslant z$ for all $z < y$, so that $\varphi(v+) \geqslant y$, which gives $\psi(y) \leqslant v$. As $v = \psi(y-) \leqslant \psi(y)$ this gives (23.2).

$-$

Pictorially, an interval where φ is constant becomes a jump of ψ, and a jump of φ becomes an interval where ψ is constant. If $\psi(b) = \psi(a+)$ then φ has a jump from a to b at $\psi(b)$. Also $h_s^* = 0$. If $\psi(a+) = \psi(a)$ there is no interval for h_s, so that (23.6) is trivially true. If $\psi(a+) > \psi(a)$ then φ is constant between $\psi(a)$ and $\psi(b)$, and $h_s = 0$, and again (23.6) is true.

Thus we can assume that $\psi(b) > \psi(a+)$. Let $\mathbf{A}$ be complete in $[\psi(a), \psi(b)]$ such that $S(\mathbf{h}; \mathbf{A}; [\psi(a), \psi(b)])$ lies in the circle centre z and radius a given $\epsilon > 0$, where z is the first integral in (23.6). Let $\mathbf{A}_1$ be complete in $[a, b]$. For each point y in $a \leqslant y < b$ we consider the intervals $[y, z)$ of $\mathscr{R}_1$. If $\psi(z) > \psi(y+)$ while $\psi(z') = \psi(y+)$ for some z' in $y < z' < z$ we omit $[y, z)$. If $\psi(z) > \psi(y+)$ while $[\psi(y+), \psi(z))$ is not in $\mathscr{R}$, we again omit $[y, z)$. Otherwise $[y, z)$ is retained; and either $\psi(z) = \psi(y+)$, which is then true for z in some right-hand neighbourhood of y, or $\psi(z') > \psi(y+)$ for each $z' > y$, and $[\psi(y+), \psi(z))$ lies in $\mathscr{R}$, so that z lies in some right-hand neighbourhood of y. Thus we obtain a family $\mathscr{R}_2$ contained in $\mathscr{R}_1$ and right-complete in $[a, b]$. Similarly we can define a left-complete family $\mathscr{L}_2$ in $[a, b]$ that is contained in $\mathscr{L}_1$; and here the situation is slightly simpler since we have (23.2). Here, if $[z, y)$ is in $\mathscr{L}_2$ then either $\psi(z+) = \psi(y)$, or $\psi(z+) < \psi(y)$ and $[\psi(z+), \psi(y))$ is in $\mathscr{L}$.

Let $\mathscr{D}$ be a division of $[a, b]$ from $\mathbf{A}_2$, and let $\mathscr{Q}$ be the collection of those intervals of $\mathscr{D}\{y, z\}$ with $\psi(z) > \psi(y+)$. Then h_s^* is zero for the other intervals of $\mathscr{D}$. In turn, let $\mathscr{Q}_1$ be the collection of intervals $[\psi(y+), \psi(z))$, for $[y, z)$ in $\mathscr{Q}$. Then $[\psi(a), \psi(b)]$, less the intervals of $\mathscr{Q}_1$, consists of a finite number of intervals of the type $[\psi(y), \psi(y+))$ for various values of y, in which φ is constant and so h_s is zero; together possibly with a finite number of points, which are discarded. We can add to $\mathscr{Q}_1$ suitable intervals from $\mathscr{L} \cup \mathscr{R}$ to make a division $\mathscr{D}_1$ of $[\psi(a), \psi(b)]$ from $\mathbf{A}$, and we have

$$(\mathscr{D}) \sum h_s^* = (\mathscr{Q}) \sum h_s^* = (\mathscr{Q}_1) \sum h_s = (\mathscr{D}_1) \sum h_s$$

and

$$S(\mathbf{h}^*; \mathbf{A}_2; [a, b])$$

lies in the circle centre z and radius ϵ. Hence (23.6).

Theorem 23.2. *If in Theorem 23.1, φ is continuous, then for $y \geqslant 0$, $\varphi(\psi(y)) = \varphi(\psi(y+)) = y$, $\psi(\varphi(y)) = \inf x$ such that $\varphi(x) = \varphi(y)$, $x \geqslant 0$.*

✓

If $x > \psi(y)$ then $\varphi(x) \geqslant y$, giving $\varphi(\psi(y)+) \geqslant y$. If $x < \psi(y)$ then $\varphi(x) < y$, $\varphi(\psi(y)-) \leqslant y$. Thus the continuity of φ gives the first result. For the second, $\psi(y+) = v$ implies that $\psi(z) \geqslant v$ for all $z > y$. Hence $v \leqslant \inf x$ such that $x \geqslant 0$, $\varphi(x) \geqslant z$. Hence if $w < v$ then $\varphi(w) < z$, and so

$$\varphi(v) = \varphi(v-) \leqslant z$$

This being true for all $z > y$, we obtain

$$y = \varphi(\psi(y)) \leqslant \varphi(\psi(y+)) = \varphi(v) \leqslant y$$

Hence the second result. The third follows from the definition of ψ.

Theorem 23.3. *Let f, g be point functions in $x \geqslant 0$, with g constant on the disjoint intervals $[u_j, v_j]$ for $j = 1, 2, \ldots$. Let*

$$\varphi(x) = x - \sum_{j=1}^{\infty} m\{(u_j, v_j)\boxed{\mathrm{n}}(0, x)\},\quad g(x) = g_1(\varphi(x))$$

should be $\cap$

New RC

where mI is the length of the interval I, with m (empty set) = 0. Let ψ be the inverse function of φ. If $f(u_j) = f(v_j)(j = 1, 2, \ldots)$, and if either integral exists, then

$$(RC) \int_{\psi(a)}^{\psi(b)} f\,dg = (RC) \int_a^b f(\psi(y))\,dg_1$$

Here, φ is continuous and monotone increasing, so that $\psi(u+) = \psi(v)(u \leqslant v)$ only if $u = v$. Thus if

$$h_s(u, v) = f(\cdot)\{g(v) - g(u)\} = f(\cdot)\{g_1(\varphi(v)) - g_1(\varphi(u)g_1(\phi(u))\}$$

$$h_s^*(y, z) = f(\psi(\cdot))\{g_1(z) - g_1(y)\}$$

then (23.3; 23.4) are satisfied. As $f(u_j) = f(v_j)(j = 1, 2, \ldots)$, and by Theorem 23.2, condition (23.5) is also satisfied, so that Theorem 23.1 gives the result, one way. Since $\varphi(\cdot -) = \varphi$

we can interchange φ and ψ, so that the result is true the other way. Colloquially, we have removed the intervals where g is constant, and have compressed the rest to a single interval.

Ex. 23.1. In Theorem 23.1 prove formally that if ψ is constant, φ has a jump, and conversely; while if ψ has a jump tden φ is constant, and conversely.

3

VARIATIONAL PROPERTIES OF THE INTEGRAL

24. The variational integral

In this chapter we give a descriptive definition of the Riemann-complete integral and define the variation and inner variation of a pair of interval functions. We begin again with Newton's integral, given in Sections 2 and 4. The property (4.1) can be written

$$|H(t)-H(x)-f(x)(t-x)|$$

$$\leqslant \begin{cases} \chi(t, x)(a < x \leqslant b, x-\delta(\epsilon, x) \leqslant t < x) \\ \chi(x, t)(a \leqslant x < b, x < t \leqslant x+\delta(\epsilon, x)) \end{cases}$$

where

$$\chi(v, w) = \epsilon(w-v) \ (a \leqslant v < w \leqslant b)$$

is a non-negative and finitely additive, and so finitely superadditive, interval function, with $\chi(a, b)$ small for small $\epsilon > 0$.

Generalizing this to pairs of interval functions and a more general χ_1, we say that the pair of finite interval functions $\mathbf{h}^*$ *is variationally equivalent to the pair* $\mathbf{h}$ *of finite interval functions in an elementary set* E if, given $\epsilon > 0$, there are an $\mathbf{A}$ complete in E, and a non-negative finitely superadditive interval function χ_1 such that

$$\chi_1(E) < \epsilon \tag{24.1}$$

$$|h_l{}^*(t, x)-h_l(t, x)| \leqslant \chi_1(t, x) \ ([t, x) \subseteq E, [t, x) \in \mathscr{L}) \tag{24.2}$$

$$|h_r{}^*(x, u)-h_r(x, u)| \leqslant \chi_1(x, u) \ ([x, u) \subseteq E, [x, u) \in \mathscr{R}) \tag{24.3}$$

If also $h_l{}^* = h_r{}^* = H$, finitely additive, then $H(E)$ is called the *variational integral of* $\mathbf{h}$ *in* E, written

$$(V)\int_E \{h_l, h_r\} = (V)\int_E \mathbf{h},$$

and we say that $\mathbf{h}$ *is variationally integrable in* E. As for Riemann-complete integration, the variational integrals of the two point functions f, g, of the single interval function h, and of $\{fh_l, fh_r\}$, are written respectively as

$$(V)\int_E f\mathrm{d}g,\ (V)\int_E h,\ (V)\int_E f\mathrm{d}\{h_l, h_r\} = (V)\int_E f\mathrm{d}\mathbf{h}$$

If the last exists, we say that f *is variationally integrable with respect to* **h** *in* E. Integrals in a set X contained in E are obtained by putting in $ch(X; x)$ as for the Riemann-complete integral.

Th. 2.5.4

NewRC [illegible]

Theorem 24.1. *If* H *is finitely additive and variationally equivalent to* **h** *in an elementary set* E, *then the Riemann-complete integral* RC(**h**; E) *exists and equals the variational integral* H(E), *which is therefore uniquely defined by* (24.1; 24.2; 24.3) *with* $\mathrm{h_l}^* = \mathrm{h_r}^* = \mathrm{H}$. *Conversely, if* RC(**h**; E) *exists, then so does the corresponding variational integral, and they are equal.*

To prove the first part let $\mathscr{D}\{v, w\}$ be a division of E from **A**. Then if S is the sum over $\mathscr{D}$ we use (24.1; 24.2; 24.3) and (21.5; 21.6) (p. 30) to obtain

$$\begin{aligned}|H(E)-S| &= |(\mathscr{D})\textstyle\sum \{H(v, w)-h_s(v, w)\}| \\ &\leqslant (\mathscr{D})\textstyle\sum |H(v, w)-h_s(v, w)| \\ &\leqslant (\mathscr{D})\textstyle\sum \chi_1(v, w) \leqslant \chi_1(E) < \epsilon\end{aligned}$$

Thus the set $S(\mathbf{h}; \mathbf{A}; E)$ lies in the circle centre $H(E)$ and radius ϵ. Hence the first part of the theorem. Conversely, let $RC(\mathbf{h}; E)$ exist. Then

$$H(v, w) = RC(\mathbf{h}; [v, w])$$

exists for all $[v, w]$ in E, and is finitely additive, by Theorem 21.1. Let us put

$$\chi_2(E_1) = \sup(\mathscr{D}_1)\textstyle\sum |h_s(v, w)-H(v, w)|$$

the supremum being for all divisions $\mathscr{D}_1\{v, w\}$ of an elementary set E_1 in E, that are formed from an $\mathbf{A}_1$ complete in E. Then by Theorem 21.2 (21.16), (p. 33), we can by suitable choice of the $\mathbf{A}_1$ suppose that χ_2 satisfies (24.1). The interval function is non-negative, and is also finitely superadditive since if $x < y < z, [x, z] \subseteq E$, a division of $[x, y]$ with a division of $[y, z]$ form a division of $[x, z]$. Taking the first two divisions so that the sums tend to the two suprema, the sum over the combined division tends to the sum of the two suprema. This proves finite superadditivity. It cannot give finite additivity since y may lie inside an interval of a division of $[x, z]$. By the method of construction of χ_2 it is clear that (24.2; 24.3) are true, and further, that if $\mathbf{A}_1$ replaces **A** then χ_2 is the least possible χ_1 having all the given properties. Thus the Riemann-complete integral is variationally equivalent to **h** in E, proving the converse.

Theorem 24.1 shows that the Riemann-complete and variational integrals are equivalent, so that properties of one integral are automatically true for the other and need not be proved again. To avoid constant repetition we omit a reference to Theorem 24.1 even when we refer to Chapter 2 for properties of the variational integral.

An equivalent definition intermediate between the Riemann-complete and variational integrals is as follows. *A finite and finitely additive interval function H is the indefinite integral in E of a pair* **h** *of interval functions if*

$$RC(\{|h_l-H|, |h_r-H|\}; E) = 0$$

The Riemann-complete integral† is wholly constructive if the complete sets **A** are constructive. However, the variational integral is descriptive, it is defined by certain of its properties. It exhibits the tight linkage between the values of the pair of interval functions and the values of the integral. Later we will see that this tight linkage enables us to prove quite easily some deep variation and differentiation properties of the integral.

The following theorem is particularly relevant when we consider previous theories that use major and minor functions. As for fixed $\mathbf{A}_1$, χ_2 is the least possible interval function for (24.2; 24.3) that is non-negative and finitely superadditive, Theorem 24.2 gives conditions under which we can have a continuous χ_2 and so continuous major and minor functions.

NewRC

Theorem 24.2. *If for fixed* x *and some*

$$\delta > 0, [\mathrm{x}-\delta, \mathrm{x}] \subseteq \mathrm{E},$$

and

$$\mathrm{j(x)} = \limsup_{\mathrm{t}\to\mathrm{x}-} \chi_2(\mathrm{t, x}) > \epsilon' > 0 \tag{24.4}$$

---

† It is as constructive as the Pollard–Getchell integral of Section 5.

where χ_2 *is the function of Theorem* 24.1, *then for a sequence* $\{t'_k\}$ *of* t *tending to* x−, *with* $[t'_k, x)$ *in* $\mathscr{R}_1$,

$$\liminf_{k\to\infty} |h_r(t'_k, x) - h_l(t'_k, x)| \geqslant \epsilon' \tag{24.5}$$

Points x *satisfying* (24.4) *for some* $\epsilon' > 0$ *are at most countable. A similar result holds for fixed* x *if for some* $\delta > 0$, $[x, x+\delta] \subseteq E$. *Here there is a sequence* $\{u'_k\}$ *tending to* x+, *with* $[x, u'_k)$ *in* $\mathscr{L}_1$, *such that*

$$\liminf_{k\to\infty} |h_l(x, u'_k) - h_r(x, u'_k)| \geqslant \epsilon' \tag{24.6}$$

If for fixed x, *and some* $\delta > 0$, $[x-\delta, x+\delta] \subseteq E$, *then at least one of* (24.5; 24.6) *or*

$$\liminf_{k\to\infty} |h_s(t'_k, u'_k) - h_l(t'_k, x) - h_r(x, u'_k)| \geqslant \epsilon' \tag{24.7}$$

holds. In (24.7), $t'_k \to x-$, $u'_k \to x+$, *while* $[t'_k, u'_k)$ *is in* $\mathscr{L}_1 \cup \mathscr{R}_1$, *with the appropriate* s.

We can obviously suppose that $E = [a, b]$. From (24.4), as χ_2 is non-negative and finitely superadditive,

$$\chi_2(t, x) > \epsilon' \; (t < x)$$

Thus over each $[t, x]$ there is a division $\mathscr{D}_2$ from A_1 such that

$$(\mathscr{D}_2) \sum |h_s - H| > \epsilon' \tag{24.8}$$

We define $\{t'_k\}$ by induction, writing $t'_1 = a$, and $[t'_{k+1}, x)$ as the interval, with right-hand end-point x, from a division $\mathscr{D}_{3k}$ of $[t'_k, x)$ satisfying (24.8). Let $\mathscr{Q}_k$ be the collection of the intervals of $\mathscr{D}_{3k}$, less the interval $[t'_{k+1}, x)$. Then for each integer m,

$$\bigcup_{k=1}^{m} \mathscr{Q}_k$$

is also a $\mathscr{Q}$ of this type, so that

$$\sum_{k=1}^{m} (\mathscr{Q}_k) \sum |h_s - H| \leqslant \chi_2(a, b) < \epsilon$$

From this, and (24.8) for the $\mathscr{D}_{3k}$,

$$\lim_{k\to\infty} (\mathscr{Q}_k) \sum |h_s - H| = 0$$

$$\liminf_{k\to\infty} |h_s(t'_k, x) - H(t'_k, x)| \geqslant \epsilon' > 0 \qquad \cancel{(24.5)}$$

By Theorem 21.2 (21.12) we can replace s by r, H by h_l, and then we have (24.5).

Now let us suppose that

$$a < t_1 < x_1 < \dots < x_{n-1} < t_n < x_n \leqslant b$$

Then as χ_2 is non-negative and finitely superadditive,

$$\sum_{k=1}^{n} \chi_2(t_k, x_k) \leqslant \chi_2(a, b) < \epsilon,$$

and

$$\sum_{k=1}^{n} j(x_k) \leqslant \epsilon, \; j(x) \geqslant 0 \; (a < x \leqslant b)$$

from which it follows that $j(x)$ can be greater than $1/m$ for not more than $m\epsilon$ points of $a < x \leqslant b$. Hence the set of points at which $j(x) > 0$, is at most countable.

Similarly for the case $x < u$. For the case $t < x < u$ a proof similar to the first part shows that if (24.5; 24.6) are false but

$$\limsup_{t,u\to x} \chi_2(t, u) > \epsilon' > 0 \; (t < x < u)$$

then

$$\liminf_{k\to\infty} |h_s(t'_k, u'_k) - H(t'_k, u'_k)| \geqslant \epsilon'$$

which gives (24.7) by the use of Theorem 21.2 (21.12; 21.13).

Ex. 24.1. Prove that χ_2 is continuous when h_l, h_r are continuous for $[t, x)$, $[x, u)$ and $[t, u)$ as $t \to x-$, $u \to x+$.

Ex. 24.2. If $\mathbf{h}_1$ is variationally equivalent to $\mathbf{h}_2$. the latter being variationally equivalent to $\mathbf{h}_3$, prove that $\mathbf{h}_1$ is variationally equivalent to $\mathbf{h}_3$. A fixed elementary set E is to be understood here. Transitivity of var: $\equiv^{ce}$. Reflexivity & symmetry also true

25. Order relation and integrability

To complete Section 20 for real interval functions, we can now give the following integrability result, which has many interesting and important applications.

Th: 2.7.5

✓ McS and my § theorem.

THEOREM 25.1. *Let* $\mathbf{h}_j$(j = 1, 2, 3) *be real and variationally integrable in an elementary set* E, *with*

$$h_{sj} \geqslant h_{s3} \; (j = 1, 2;\ s = l, r), \quad h_{s4} = \max(h_{s1}, h_{s2}) \; (s = l, r) \tag{25.1}$$

Then $\mathbf{h}_4$ *is also variationally integrable in* E.

By Theorem 19.1 the variational integral is a distributive functional. Thus we can subtract h_{s3} from the other three pairs of interval functions in (25.1) and so assume that

$$h_{sj} \geqslant 0 \; (j = 1,\ 2;\ s = l, r),\quad h_{s4} = \max(h_{s1}, h_{s2}) \; (s = l, \mathrm{r}) \tag{25.2}$$

Let H_j be the variational integral of $\mathbf{h}_j (j = 1, 2)$, and put

$$H_3 = \max(H_1, H_2)$$

Then given $\epsilon > 0$, there are χ_j, non-negative and finitely superadditive and $\mathbf{A}_j$ complete in $E (j = 1, 2)$, such that

$$\chi_j(E) < \tfrac{1}{2}\epsilon \; (j = 1, 2) \tag{25.3}$$

$$|h_{sj} - H_j| < \chi_j \tag{25.4}$$

($s = l, r,$ *for intervals in* $\mathscr{L}_j$, $\mathscr{R}_j$, *respectively; and* $j = 1, 2$).

Putting $\mathbf{A}_3 = \mathbf{A}_1.\mathbf{A}_2$, and taking the intervals from the appropriate family $\mathscr{L}_3$ or $\mathscr{R}_3$, (25.4) gives

$$h_{sj} < H_j + \chi_j \leqslant H_3 + \chi_1 + \chi_2 \; (j = 1, 2)$$

$$h_{s4} < H_3 + \chi_1 + \chi_2 \tag{25.5}$$

Also from (25.4),

$$H_j - \chi_1 - \chi_2 \leqslant H_j - \chi_j < h_{sj} \leqslant h_{s4} \; (j = 1, 2)$$

$$H_3 - \chi_1 - \chi_2 < h_{s4} \tag{25.6}$$

From (25.3; 25.5; 25.6), $\mathbf{h}_4$ is variationally equivalent to H_3. However, H_3 need not be finitely additive, so that it is not necessarily the variational integral. We have to integrate H_3, and we note that

$$H_3(u, v) + H_3(v, w) \geqslant H_j(u, v) + H_j(v, w)$$

$$= H_j(u, w) \; (j = 1, 2;\ u < v < w, [u, w] \subseteq E),$$

$$H_3(u, w) \leqslant H_3(u,v) + H_3(v, w) \; (u < v < w, [u, w] \subseteq E) \tag{25.7}$$

and H_3 is finitely subadditive in E. Also, successively, we have

$$h_{sj} \geqslant 0 \; (j = 1, 2;\ s = l, r),\quad H_j \geqslant 0 \; (j = 1, 2),$$

$$H_j \leqslant H_1 + H_2 \; (j = 1, 2),\quad H_3 \leqslant H_1 + H_2 \tag{25.8}$$

and $H_1 + H_2$ is finitely additive. By (25.7; 25.8), H_3 satisfies all the conditions of Theorem 22.1, first part, and so has a Riemann-complete integral, H_4 say, that is variationally equivalent to H_3. In turn this is variationally equivalent to $\mathbf{h}_4$. Thus by Ex. 24.2 there exists

$$(V)\int_E \mathbf{h}_4 = H_4(E)$$

i.e.

$$(V)\int_E \{\max_{j=1,2} h_{lj}, \max_{j=1,2} h_{rj}\}$$

$$= (V)\int_E \max_{j=1,2}\left\{(V)\int \{h_{lj}, h_{rj}\}\right\} \tag{25.9}$$

proving the theorem.

From Theorem 25.1 we can deduce the following important results concerning absolute integrability of pairs of real interval functions. True for h complex, see Th. 5 of [illegible] of interv

Theorem 25.2. *Let* $\mathbf{h}$ *be real and variationally integrable in an elementary set* E, *and let there be a set* $\mathbf{A}_4$ *complete in* E, *and a number* M, *such that every division* $\mathscr{D}$ *of* E *from* $\mathbf{A}_4$ *gives a sum* Th. 2.7.3

Done

$$(\mathscr{D})\sum |\mathbf{h}_s| \leqslant \mathrm{M}$$

Then $\{|\mathbf{h}_l|, |\mathbf{h}_r|\} \equiv |\mathbf{h}|$ *is variationally integrable in* E.

Let us put

$$h_5(I) = \sup(\mathscr{D}_1)\sum |h_s|$$

over all divisions $\mathscr{D}_1$ of I from $\mathbf{A}_4$. By adding a division from $\mathbf{A}_4$ over the elementary set formed by closing the intervals of $E \cap CI$, we see that

$$0 \leqslant h_5(I) \leqslant M (I \subseteq E)$$

so that h_5 is finite. It is clearly finitely superadditive with $S(h_5; \mathbf{A}_4; E)$ bounded below by 0, so that, by Theorem 22.1, h_5 is Riemann-complete integrable. Thus Theorem 25.1 gives the result, since

$$h_s \geqslant -h_5,\quad -h_s \geqslant -h_5,\quad |h_s| = \max(h_s, -h_s)$$

Theorem 25.3. *Let* $\{\mathrm{fh}_l, \mathrm{fh}_r\}$ *be real and variationally integrable in an elementary set* E, *and let* $|\mathrm{f}| \leqslant \mathrm{g}$, *where*

$\{g|h_l|, g|h_r|\}$ *is variationally integrable in* E.
Then $\{|f||h_l|, |f||h_r|\}$ *is variationally integrable in* E.
We can apply Theorem 25.1 since

$$fh_s \geqslant -g|h_s|,\ -fh_s \geqslant -g|h_s|,\ \max(fh_s, -fh_s) = |f|.|h_s|$$

Theorem 25.4. Let $\mathbf{h}$, $|\mathbf{h}|$, $\{fh_l, fh_r\}$, $\{|f||h_l|, |f||h_r|\}$ *be real and variationally integrable in* E. *Then if* α *is a real constant*

$$\{\max(fh_l, \alpha h_l), \max(fh_r, \alpha h_r)\}$$

and

$$\{\min(fh_l, \alpha h_l), \min(fh_r, \alpha h_r)\}$$

are also variationally integrable in E.
For by the distributivity of the integral

$$\{(|f|+|\alpha|)|h_l|, (|f|+|\alpha|)|h_r|\}\ \{\alpha h_l, \alpha h_r\}$$

are also variationally integrable in E, while

$$(-|f|-|\alpha|)|h_s| \leqslant fh_s,\ (-|f|-|\alpha|)|h_s| \leqslant \alpha h_s\ (s = l, r)$$

Thus Theorem 25.1 gives the result for the max, while the min follows from

$$\min(fh_s, \alpha h_s) = \alpha h_s + fh_s - \max(fh_s, \alpha h_s)\ (s = l, r)\ .$$

Theorem 25.4 is usually considered for the case $h_s \geqslant 0$, $h_l = h_r$, these being finitely additive.

26. Variation

NV

The property of variational equivalence of $\mathbf{h}^*$ with $\mathbf{h}$ is by (24.1) for the non-negative finitely superadditive χ_1, equivalent to a statement that in some sense, $\{h_l^* - h_l, h_r^* - h_r\}$ is of variation zero in E. To generalize this property, a pair $\mathbf{h}$ of interval functions is said to be *of bounded variation* (VB^*) *in an elementary set E* if there are an $\mathbf{A}$ complete in E and a non-negative finitely superadditive interval function χ such that $\chi(E)$ is finite and

$$|h_l(t, x)| \leqslant \chi(t, x)\ \ ([t, x) \in \mathscr{L}, [t, x) \subseteq E) \tag{26.1}$$

$$|h_r(x, u)| \leqslant \chi(x, u)\ ([x, u) \in \mathscr{R}, [x, u) \subseteq E) \tag{26.2}$$

The variation of $\mathbf{h}$ *in* E *is*

$$V(\mathbf{h}; E) = \inf \chi(E) \tag{26.3}$$

for all such χ. If there is no such χ we write the right-hand side symbolically as $+\infty$. If the infimum is 0 we say that $\mathbf{h}$ *is of variation zero in E*. Thus $\mathbf{h}$ and $\mathbf{h}^*$ are variationally equivalent in E when $\{h_l - h_l^*, h_r - h_r^*\}$ is of variation zero in E. Further we can now replace in Theorem 25.2 the condition that the sums over divisions, of the moduli, should be bounded, by the condition that $\mathbf{h}$ is of bounded variation. Similar definitions can be found for a point function g by putting

$$h_l(t, u) = h_r(t, u) = g(u) - g(t)\ (\text{all } [t, u] \subseteq E)$$

More generally, let X be a set of x, not necessarily lying entirely in E. Then the definitions that $\mathbf{h}$ *is of bounded variation in X, relative to E*, and of the *variation $V(\mathbf{h}; E; X)$ in X, relative to E*, are obtained on adding ($x \in X$) to the conditions in (26.1; 26.2). Thus we can now define the properties that $\mathbf{h}$ *is of variation zero in X, relative to E, and that* $\mathbf{h}$, $\mathbf{h}^*$ *are variationally equivalent in X, relative to E*, in a similar way.†

The definition of (26.3) can be split up into two parts, putting

$$V^*(\mathbf{h}, E; \mathbf{A}) = \inf \chi(E)$$

for fixed $\mathbf{A}$ and all such χ. Then $V(\mathbf{h}; E)$ is the limit of V^* as we take smaller and smaller sets $\mathbf{A}$. For if $\mathbf{A}_1$ is also complete in E, with the product $\mathbf{A}.\mathbf{A}_1 = \mathbf{A}_1$, then a χ suitable for $\mathbf{A}$ will also be suitable for $\mathbf{A}_1$, which gives

$$V^*(\mathbf{h}; E; \mathbf{A}) \geqslant V^*(\mathbf{h}; E; \mathbf{A}_1)$$

The converse need not be true when $\mathbf{A}$, $\mathbf{A}_1$ are unequal.

Ex. 26.1. $V(|\mathbf{h}|; E) = V(\mathbf{h}; E)$.

Ex. 26.2. $V(\mathbf{h}; E; X) = V(\{ch(X)h_l, ch(X)h_r\}; E)$.

Ex. 26.3. If $W = \sup|z|$ (*all* $z \in P(\mathbf{h}; \mathbf{A}_2; E)$), show that

$$W \leqslant V^*(\mathbf{h}; E; \mathbf{A}_2) \leqslant 4W$$

(Use the method of proof of Theorem 21.2 (21.16).) Thus $\mathbf{h}$ is of bounded variation if and only if P is a bounded set for some $\mathbf{A}_2$; and $\mathbf{h}$ is of variation zero if and only if, given $\epsilon > 0$, there is an $\mathbf{A}_2$ such that the corresponding P lies in the circle centre 0 and radius ϵ. These are the reasons for the name, *variation set*, for P. When h_s take values that are not (real or) complex numbers, the variation is often defined by using P.

---

† In this new theory the variation in X, relative to E, takes the place of Lebesgue outer measure.

Ex. 26.4. $V^*(\mathbf{h}; E; \mathbf{A}) = \sup(\mathscr{D}) \Sigma|h_s|$ for all divisions $\mathscr{D}$ of E from $\mathbf{A}$.

Ex. 26.5. If h_l, h_r are complex, with real parts h_{l1}, h_{r1}, respectively, and imaginary parts h_{l2}, h_{r2}, respectively, prove that

$$V(\{h_{lj}, h_{rj}\}; E; X) \leqslant V(\{h_l, h_r\}; E; X)(j = 1, 2)$$

Ex. 26.6. If $X \subseteq X_1, E \subseteq E_1$, prove that

$$V(\mathbf{h}; E; X) \leqslant V(\mathbf{h}; E_1; X_1)$$

Ex. 26.7. Let α, β be constants, and let $\mathbf{h}, \mathbf{h}^*$ be two pairs of interval functions. Prove that

$$V(\{\alpha h_l, \alpha h_r\}; E; X) = |\alpha| V(\{h_l, h_r\}; E; X),$$

$$V(\{\alpha h_l + \beta h_l^*, \alpha h_r + \beta h_r^*\}; E; X)$$
$$\leqslant |\alpha| V(\{h_l, h_r\}; E; X) + |\beta| V(\{h_l^*, h_r^*\}; E; X)$$

27. Inner variation

NV

A concept needed when we consider differentiation, is that of the *inner variation* of a pair $\mathbf{h}$ of interval functions. To define it we first say that a set $\mathscr{C}$ of intervals $[u, v)$ is an *inner cover* of an elementary set E if to each point x of E there corresponds at least one interval from $\mathscr{C}$ and in E with x as its associated point. Given $\mathbf{A}$ complete in E, let there be an inner cover $\mathscr{C}$ of E formed from the intervals of $\mathscr{L}$ and $\mathscr{R}$, and a non-negative finitely superadditive interval function χ, such that $\chi(E)$ is finite and

$$|h_l(t, x)| \leqslant \chi(t, x) \ ([t, x) \in \mathscr{C} \cap \mathscr{L}, [t, x) \subseteq E) \qquad (27.1)$$

$$|h_r(x, u)| \leqslant \chi(x, u) \ ([x, u) \in \mathscr{C} \cap \mathscr{R}, [x, u) \subseteq E) \qquad (27.2)$$

Fixing $\mathbf{A}$, we put

$$IV^*(\mathbf{h}; E; \mathbf{A}) = \inf \chi(E)$$

for all such χ. If no such χ exists we put $IV^* = +\infty$. Now let $\mathbf{A}_1$ be complete in E, with $\mathbf{A}_1 = \mathbf{A}.\mathbf{A}_1$. Then a $\mathscr{C}$ from $\mathbf{A}_1$ is also from $\mathbf{A}$ and hence

$$IV^*(\mathbf{h}; E; \mathbf{A}) \cancel{\leqslant} IV^*(\mathbf{h}; E; \mathbf{A}_1)$$

The converse need not be true, as the chosen intervals of

$\geqslant$

[illegible] again p. 136 similarly for IV = 0

$\mathscr{L}, \mathscr{R}$, need not lie in $\mathscr{L}_1, \mathscr{R}_1$, respectively. Then $\mathbf{h}$ *is of bounded inner variation in an elementary set E if the inner variation*

$$IV(\mathbf{h}; E) = \overset{\text{inf}}{\cancel{\sup}}\, IV^*(\mathbf{h}; E; \mathbf{A}) \qquad (27.3)$$

in E is finite, the supremum being taken for all $\mathbf{A}$ complete in E. Note that the supremum is the limit of IV^* as we take smaller and smaller $\mathbf{A}$. Also, if X is a set not necessarily lying entirely in E, we define

$$IV^*(\mathbf{h}; E; X; \mathbf{A}) = IV^*(\{ch(X)h_l, ch(X)h_r\}; E; \mathbf{A})$$
$$IV(\mathbf{h}; E; X) = IV(\{ch(X)h_l, ch(X)h_r\}; E)$$

the second line giving the *inner variation in X relative to E*. If this is finite we say that $\mathbf{h}$ *is of bounded inner variation in X relative to E*. Further, $\mathbf{h}$ *is of inner variation zero in E, or in X relative to E*, if the corresponding inner variation is 0.

Ex. 27.1. Prove that

$$IV(\mathbf{h}; E; X) \leqslant V(\mathbf{h}; E; X)$$

Ex 27.2. Prove results for IV corresponding to Ex. 26.1; ~~26.5;~~ 26.6; 26.7 (first part).

(for $E = E_1$)

28. Properties of the variation and inner variation

THEOREM 28.1. *If*

✓

NV

$$\mathrm{X} = \bigcup_{j=1}^{\infty} \mathrm{X}_{1j}$$

then Th. 2.2.1 4th blk (not IV)

$$\mathrm{V}(\mathbf{h}; \mathrm{E}; \mathrm{X}) \leqslant \sum_{j=1}^{\infty} \mathrm{V}(\mathbf{h}; \mathrm{E}; \mathrm{X}_{1j})$$

3rd blk

and similarly for IV. In Th. 15.6 (15.2) p. 137 when the IV = 0

We can obviously assume that the infinite series is a finite number, or else there is nothing to prove. By Ex.26.6 we can also assume that the X_{1j} are mutually disjoint. Given $\epsilon > 0$ and an integer j, there are $\mathbf{A}_{1j}$ complete in E, and a non-negative finitely superadditive interval function χ_{1j}, that satisfy (26.1; 26.2) with x in X_{1j} and

$$\chi_{1j}(E) < V(\mathbf{h}; E; X_{1j}) + 2^{-j}\epsilon \qquad (28.1)$$

Let the $\mathbf{A}_6(x)$ of Ex. 16.9 be $\mathbf{A}_{1j}$ for x in $X_{1j} (j = 1, 2, \ldots)$, and be arbitrary for x not in X. We then denote the resulting $\mathbf{A}_7$ of Ex.16.9 by $\mathbf{A}$ here. Further, we put

$$\chi = \sum_{j=1}^{\infty} \chi_{1j}$$

Then by (28.1) and the definition of variation,

$$V(\mathbf{h}; E; X) \leqslant \chi(E) = \sum_{j=1}^{\infty} \chi_{1j}(E) < \sum_{j=1}^{\infty} V(\mathbf{h}; E; X_{1j}) + \epsilon$$

As $\epsilon > 0$ is arbitrary we prove the theorem. Similarly for IV^* and so IV.

THEOREM 28.2. *If* $\mathrm{E}_1, \mathrm{E}_2$ *are non-overlapping elementary sets,*

$\leq$ in Th. 25.6
in Th. 27.7

(NV) ✓ $$\mathrm{V}(\mathbf{h}; \mathrm{E}_1; \mathrm{X}) + \mathrm{V}(\mathbf{h}; \mathrm{E}_2; \mathrm{X}) = \mathrm{V}(\mathbf{h}; \mathrm{E}_1 \cup \mathrm{E}_2; \mathrm{X}) \tag{28.2}$$

$$\mathrm{IV}(\mathbf{h}; \mathrm{E}_1; \mathrm{X}) + \mathrm{IV}(\mathbf{h}; \mathrm{E}_2; \mathrm{X}) \geqslant \mathrm{IV}(\mathbf{h}; \mathrm{E}_1 \cup \mathrm{E}_2; \mathrm{X}) \tag{28.3}$$

with equality in (28.3) *if* $\mathrm{E}_1 \cap \mathrm{E}_2 \cap \mathrm{X}$ *is empty.*
From Ex. 26.6

$$V(\mathbf{h}; E_j; X) \leqslant V(\mathbf{h}; E_1 \cup E_2; X) (j = 1, 2)$$

Thus we can suppose that the left-hand side of (28.2) is finite, or else the right-hand side is infinite and there is nothing to prove. The essence of the proof is that the intervals in E_1 can vary independently of the intervals in E_2. Given $\epsilon > 0$, we take x in X, and χ_{2j} for χ, E_j for E, in (26.1; 26.2), such that

$$\chi_{2j}(E_j) < V(\mathbf{h}; E_j; X) + \epsilon \ (j = 1, 2)$$

with χ_{21} zero in E_2, χ_{22} zero in E_1..

Then we can put $\chi_{21} + \chi_{22}$ for χ, $E_1 \cup E_2$ for E, in (26.1; 26.2)

$$V(\mathbf{h}; E_1 \cup E_2; X) \leqslant \chi_{21}(E_1) + \chi_{22}(E_2)$$

$$< V(\mathbf{h}; E_1; X) + V(\mathbf{h}; E_2; X) + 2\epsilon$$

Smaller

$$V(\mathbf{h}; E_1 \cup E_2; X) \leqslant V(\mathbf{h}; E_1; X) + V(\mathbf{h}; E_2; X) \tag{28.4}$$

The corresponding result for IV^*, and (28.3), follow similarly.

We now replace χ by χ_3, E by $E_1 \cup E_2$, in (26.1; 26.2) with x in X and $\chi_3(E_1 \cup E_2) < V(\mathbf{h}; E_1 \cup E_2; X) + \epsilon$. Then χ_3 will be suitable in E_1, E_2 separately, and

$$V(\mathbf{h}; E_1; X) + V(\mathbf{h}; E_2; X) \leqslant \chi_3(E_1) + \chi_3(E_2)$$

$$\leqslant \chi_3(E_1 \cup E_2) < V(\mathbf{h}; E_1 \cup E_2; X) + \epsilon$$

As $\epsilon > 0$ is arbitrary we obtain the opposite inequality to (28.4), and so equality. There need not be equality in (28.3), for if x lies in $X_1 \cap E_1 \cap E_2$, the intervals of (27.1; 27.2), with $E_1 \cup E_2$ for E, and with associated point x, might all lie in E_1, or all in E_2. However, when $X \cap E_1 \cap E_2$ is empty this is impossible, and the proof of equality in (28.3) goes through as before, provided that every interval of the left-complete and right-complete sets used lies entirely within either E_1 or E_2. This restriction on complete sets for IV^* makes no difference to IV, so that the proof is complete.

THEOREM 28.3. *Let* $\{\mathrm{I}_j\}$ *be a finite or infinite sequence of disjoint intervals with union* J. *If each two intervals with a common end-point are open at that end, then*

(NV)
As Ex. 5.3

$$\mathrm{V}(\mathbf{h}; \mathrm{E}; \mathrm{X} \cap \mathrm{J}) = \sum_{j=1}^{\infty} \mathrm{V}(\mathbf{h}; \mathrm{E}; \mathrm{X} \cap \mathrm{I}_j)$$

Consider the intervals $I_1, \ldots, I_n$. By the remark about common end-points there is a collection of intervals $J_1, \ldots, J_n$, open on the right, and mutually disjoint, such that each point of I_j is inside $J_j (j = 1, \ldots, n)$. Given $\epsilon > 0$, let $\mathbf{A}_2$ be complete in E, and χ_4 be non-negative and finitely superadditive, such that

$$\chi_4(E) < V\left(\mathbf{h}; E; X \cap \bigcup_{j=1}^{n} I_j\right) + \epsilon \tag{28.5}$$

$$|h_s| \leqslant \chi_4 \tag{28.6}$$

for the intervals of $\mathscr{L}_2 \cup \mathscr{R}_2$, with associated points in

$$X \cap \bigcup_{j=1}^{n} I_j$$

We now discard every interval in $\mathscr{L}_2 \cup \mathscr{R}_2$ that does not lie entirely within either one of the J_j, or in

$$E \cap \bigcap_{j=1}^{n} CJ_j \qquad \cancel{(28.5;\ 26.6)}$$

and we are still left with a complete set. This shows that from (28.5; 28.6),

$$\sum_{j=1}^{n} V(\mathbf{h}; E; X \cap I_j) \leqslant \sum_{j=1}^{n} \chi_4(J_j)$$

$$\leqslant \chi_4(E) < V(\mathbf{h}; E; X \cap \bigcup_{j=1}^{n} I_j) + \epsilon$$

By Ex. 26.6 and Theorem 28.1, the last is

$$\leqslant V(\mathbf{h}; E; X \cap J) + \epsilon \leqslant \sum_{j=1}^{\infty} V(\mathbf{h}; E; X \cap I_j) + \epsilon$$

As $\epsilon > 0$ is arbitrary we have the result, even if one side is infinite.

Theorem 28.4. *If* $\mathrm{V}(\mathbf{h}; \mathrm{E}; \mathrm{X})$ *is finite, then given* $\epsilon > 0$, *there are* $\mathbf{A}_3$ *complete in* E, *and a non-negative finitely superadditive interval function* χ_5, *such that*

Th. 2.7.8(2.7.2IA)

$$\chi_5(\mathrm{E}) < \epsilon, \tag{28.7}$$

$$|\mathrm{h}_1(\mathrm{t}, \mathrm{x})| \leqslant \mathrm{V}(\mathbf{h}; [\mathrm{t}, \mathrm{x}); \mathrm{X}) + \chi_5(\mathrm{t}, \mathrm{x})$$

$$(\mathrm{x} \in \mathrm{X}, [\mathrm{t}, \mathrm{x}) \in \mathscr{L}_3, [\mathrm{t}, \mathrm{x}) \subseteq \mathrm{E}) \tag{28.8}$$

$$|\mathrm{h}_r(\mathrm{x}, \mathrm{u})| \leqslant \mathrm{V}(\mathbf{h}; [\mathrm{x}, \mathrm{u}); \mathrm{X}) + \chi_5(\mathrm{x}, \mathrm{u})$$

$$(\mathrm{x} \in \mathrm{X}, [\mathrm{x}, \mathrm{u}) \in \mathscr{R}_3, [\mathrm{x}, \mathrm{u}) \subseteq \mathrm{E}) \tag{28.9}$$

We choose $\mathbf{A}_3$, χ_6 satisfying (26.1; 26.2) with x in X, and such that

$$\chi_6(E) < V(\mathbf{h}; E; X) + \epsilon \tag{28.10}$$

By Theorem 28.2 the variation is finitely additive for E_1 in E, so that

$$\chi_5(E_1) = \chi_6(E_1) - V(\mathbf{h}; E_1; X)$$

is finitely superadditive. By (28.10) we have (28.7), and by (26.1; 26.2) we have (28.8; 28.9). Finally, the same χ_6 will be

suitable for each E_1 in E, so that χ_5 is non-negative. Hence the theorem.

Theorem 28.5. *Given* $\mathbf{h}$,VB* *in* E, $\epsilon > 0$, *and* $\mathbf{A}_4$ *complete in* E, *then there is a division* $\mathscr{D}$ *of* E *from* $\mathbf{A}_4$ *such that every collection* $\mathscr{Q}_1$ *of intervals from* $\mathscr{D}$, *with a corresponding collection* $\mathscr{Q}$ *of the closures of the intervals of* $\mathscr{Q}_1$, *satisfies*

$$(\mathscr{Q}_1) \sum |h_s| \geqslant V(\mathbf{h}; \bigcup_{\mathscr{Q}} I) - 2\epsilon \tag{28.11}$$

Further, if $[\mathrm{t}, \mathrm{x}) \subseteq \mathrm{E}$ *for some* $\mathrm{t} < \mathrm{x}$ *then*

Ex5.10 pp 73-74

$$\lim_{t \to x-} \sup |h_l(t, x)| = \lim_{t \to x-} V(\mathbf{h}; [t, x)) \tag{28.12}$$

and similarly for $\mathrm{u} \to \mathrm{x}+$, *if* $[\mathrm{x}, \mathrm{u}) \subseteq \mathrm{E}$ *for some* $\mathrm{u} > \mathrm{x}$.

The least χ_7 to satisfy (26.1; 26.2) with $\mathbf{A}_4$ replacing $\mathbf{A}$, and E_1 replacing E, is the interval function

$$\chi_7(E_1) = \sup(\mathscr{D}_1) \sum |h_s|$$

for all divisions $\mathscr{D}_1$ of E_1 from $\mathbf{A}_4$. It follows from the definition of the variation that there is a division $\mathscr{D}$ of E and from $\mathbf{A}_4$, such that

$$(\mathscr{D}) \sum |h_s| > V(\mathbf{h}; E) - \epsilon \tag{28.13}$$

Replacing $\mathbf{A}_4$ by $\mathbf{A}_3.\mathbf{A}_4$, we can also assume that the intervals of $\mathscr{D}$ satisfy (28.7; 28.8; 28.9) of Theorem 28.4. From these, (28.13), and Theorem 28.2,

$$\begin{aligned}(\mathscr{Q}_1) \sum |\mathrm{h}_s| &= (\mathscr{D}) \sum |h_s| - (\mathscr{D} - \mathscr{Q}_1) \sum |h_s| \\ &\geqslant V(\mathbf{h}; E) - \epsilon - (\mathscr{D} - \mathscr{Q}_1) \sum [V(\mathbf{h}; I) + \chi_5(I)] \\ &\geqslant V(\mathbf{h}; \bigcup_{\mathscr{Q}} I) - 2\epsilon\end{aligned}$$

giving (28.11).

For (28.12) we can also assume that no interval $[t, x)$ lies in $\mathscr{R}_4$, for we can discard such intervals (for the fixed x) without discarding the completeness of $\mathbf{A}_4$. Then (28.11) gives

$$\sup\{|h_l(t, x)| - V(\mathbf{h}; [t, x])\} \geqslant -2\epsilon \ ([t, x) \in \mathscr{L}_4)$$

In this we can take $\epsilon \to 0$, if at the same time we make the defining interval at x of $\mathscr{L}_4$ shrink up to x. Thus we obtain

$$\lim_{t \to x-} \sup[|h_l(t, x)| - V(\mathbf{h}; [t, x])] \geqslant 0 \tag{28.14}$$

The variation is non-negative, and is finitely additive by Theorem 28.2. Hence it is monotone decreasing as $t \to x-$, so that by Theorem 13.2 it tends to a non-negative limit. Thus (28.14) gives

$$\limsup |h_l(t, x)| \geqslant \lim V(\mathbf{h}; [t, x])$$

Theorem 28.4 gives the opposite inequality, so that we have (28.12).

Similarly for intervals $[x, u)$.

Th, 28.1

THEOREM 28.6. *Suppose that*

NV?

$$X_{21} \subseteq X_{22} \subseteq \ldots \subseteq X_{2j} \subseteq \ldots, \lim_{j\to\infty} X_{2j} = X_3$$

Then

V( ) ⟶

$$\lim_{j\to\infty} V(\mathbf{h}; E; X_{2j}) = V(\mathbf{h}; E; X_3) \qquad (28.15)$$

If $V(\mathbf{h}; E; X)$ *is finite for* $X = X_{41}$, *and finitely additive over differences of sets* X_{4j}, *where*

$$X_{41} \supseteq X_{42} \supseteq \ldots \supseteq X_{4j} \supseteq \ldots, \lim_{j\to\infty} X_{4j} = X_5$$

then we have a similar result.†

In Theorem 28.4 we take X_{2j}, $\epsilon.2^{-j}$, A_{5j}, χ_{8j}, for X, ϵ, A_3, χ_5, putting

$$K_j(E_1) = V(\mathbf{h}; E_1; X_{2j}),\ K(E_1) = \lim_{j\to\infty} K_j(E_1)$$

which we can assume finite for each elementary set E_1 in E, or else the theorem is true by Ex. 26.6. Then for I in $\mathscr{L}_{5j} \cup \mathscr{R}_{5j}$, with associated point in X_{2j}

$$|h_s| \leqslant K_j + \chi_{8j} \leqslant K + \sum_{j=1}^{\infty} \chi_{8j} \qquad (28.16)$$

We now take the $A_6(x)$ of Ex. 16.9 (p.24) to be A_{5j} when x is in $X_{2j} \cap CX_{2,j-1}$, and to be arbitrary when x is in CX_3, and we can construct the A_7 as usual, where here we denote it by A_6. The K_j, and so K, are finitely additive, so that

$$K + \sum_{j=1}^{\infty} \chi_{8j}$$

---

† The corresponding theorem for outer measure has to assume that the outer measure is regular. Here the variation is not always regular.

is finitely superadditive. Thus by (28.16) and the definition of variation,

$$V(\mathbf{h}; E; X_3) \leqslant K(E) + \sum_{j=1}^{\infty} \chi_{8j}(E) < K(E) + \epsilon$$

$$V(\mathbf{h}; E; X_3) \leqslant K(E) = \lim_{j\to\infty} V(\mathbf{h}; E; X_{2j})$$

The reverse inequality is true by Ex. 26.6, (p.48) so that we have the first case of (28.15). For the second, $X_{41} \cap CX_{4j}$ is monotone increasing in j, so that by the first part, Theorem 28.1, and the special finite additivity, in X_{4j}, $X_{41} \cap CX_{4j}$,

$$V(\mathbf{h}; E; X_5) \geqslant V(\mathbf{h}; E; X_{41}) - V(\mathbf{h}; E; X_{41} \cap CX_5)$$

$$= V(\mathbf{h}; E; X_{41}) - \lim_{j\to\infty} V(\mathbf{h}; E; X_{41} \cap CX_{4j})$$

$$= \lim_{j\to\infty} V(\mathbf{h}; E; X_{4j})$$

$$V(\mathbf{h}; E; X_5) \geqslant \lim_{j\to\infty} V(\mathbf{h}; E; X_{4j})$$

The reverse inequality is true by Ex.26.6, so that we have proved the second result of the theorem.

Ex. 28.1. If

$$X = \bigcap_{j=1}^{\infty} X_{1j} \qquad \textcircled{\cup}$$

and if $\mathbf{h}$ is of variation (inner variation) zero in each X_{1j} relative to E, show that it is the same† in X, relative to E.

Ex. 28.2. For $a < b$, $[a, b] \subseteq E$, $x = a$ and $x = b$, let

$$\lim_{\delta\to 0} V(\mathbf{h}; E; [x-\delta, x+\delta]) = 0$$

ital

Prove that for $I = (a, b), [a, b), (a, b]$, and $[a, b]$,

$$V(\mathbf{h}; [a, b]) = V(\mathbf{h}; E; I)$$

(For $[a, b] \subseteq [a-\delta, a+\delta] \cup (a, b) \cup [b-\delta, b+\delta]$, and

$$V(\mathbf{h}; [a, b]) = V(\mathbf{h}; [a, b]; [a, b]))$$

Ex. 28.3. By using Theorem 28.1 with $h_s(I)$ as the length of the interval I, show that $[0, 1]$ is not a countable set.

[illegible]

† This corresponds to the result in Lebesgue theory that a countable union of sets of measure zero is also of measure zero.

Ex. 28.4. By using the interval function of Ex. 31.2, p.64, with $E = X =$ show that there is no set $X_1 \supseteq X$ for which there exists $(v)\int_E dh(X_1)$

29. Functions of generalized bounded variation

A less restrictive condition than that of bounded variation, is as follows. The pair **h** of interval functions is of *generalized bounded variation* (*VBG**) *in a set X, relative to an elementary set E*, if *X* is the union of an at most countable number of sets in each of which **h** is of bounded variation relative to *E*. If the sum of the variations is convergent, then by Theorem 28.1 we have nothing new.

NV

Theorem 29.1. **h** *is* VBG* *in* X, *relative to* E, *if and only if there are* **A** *complete in* E, *a non-negative finitely superadditive interval function* χ, *with* $\chi(E)$ *finite, and a point function* $k(x) > 0$, *such that*

✓

$$|\mathbf{h}_l(t, x)| \leqslant k(x)\chi(t, x) \quad (x \in X, [t, x) \in \mathscr{L}, [t, x) \subseteq E) \tag{29.1}$$

$$|\mathbf{h}_r(x, u)| \leqslant k(x)\chi(x, u) \quad (x \in X, [x, u) \in \mathscr{R}, [x, u) \subseteq E) \tag{29.2}$$

If **h** is *VBG** in *X* relative to *E*, there are sets X_{1j} with union *X*, in each of which **h** is *VB** relative to *E*. By Ex. 26.6 we can assume the X_{1j} mutually disjoint, and then as in Ex. 16.9 we can obtain **A** by combining the separate complete sets for the sets X_{1j} of *x*. If the corresponding χ are χ_{1j} we can put

p.48,

24,

$$\chi = \sum_{j=1}^{\infty} 2^{-j}\chi_{1j}/\chi_{1j}(E),\ k(x) = 2^j\chi_{1j}(E) \quad (x \in X_{1j}, j = 1, 2, \ldots)$$

and then we have (29.1; 29.2).

Conversely, if (29.1; 29.2) are true we can take X_{2j} as the set where

$$j-1 \leqslant k(x) < j, \text{ with } \chi_{2j} = j\chi,$$

using the **A** of (29.1; 29.2), and we see that **h** is *VB** in each X_{2j}, relative to *E*, and so *VBG** in *X*, relative to *E*.

We could similarly define the property of being of *generalized bounded inner variation*, and we could obtain a similar theorem.

30. The variation of variationally equivalent functions

This section begins with an important theorem that makes many proofs very simple.

NV

Theorem 30.1. *If* **h** *is variationally equivalent to* **h*** *in* X, *relative to* E, *and if* X_1 *is a set in* X, *then*

✓

$$V(\mathbf{h}; E; X_1) = V(\mathbf{h}^*; E; X_1) \tag{30.1}$$

Ex.2.1.7

$$IV(\mathbf{h}; E; X_1) = IV(\mathbf{h}^*; E; X_1) \tag{30.2}$$

even if a V *or an* IV *is infinite. Again, let* f *be a point function, and let* **h** *be variationally equivalent to* $\{f\mathbf{h}_l^*, f\mathbf{h}_r^*\}$ *in* X, *relative to* E. *If* $|f| \leqslant C$ *in* X_2 *contained in* X, *then*

For V see Ex.2.18

$$V(\mathbf{h}; E; X_2) \leqslant C.V(\mathbf{h}^*; E; X_2) \tag{30.3}$$

$$IV(\mathbf{h}; E; X_2) \leqslant C.IV(\mathbf{h}^*; E; X_2) \tag{30.4}$$

If, however, $|f| \geqslant C$ *in* X_3 *contained in* X, *then*†

$$V(\mathbf{h}; E; X_3) \geqslant C.V(\mathbf{h}^*; E; X_3) \tag{30.5}$$

$$IV(\mathbf{h}; E; X_3) \geqslant C.IV(\mathbf{h}^*; E; X_3) \tag{30.6}$$

To prove (30.1) we note that, given $\epsilon > 0$, there are **A** complete in *E*, and a non-negative finitely superadditive interval function χ satisfying

$$\chi(E) < \epsilon \tag{30.7}$$

$$|h_l(t, x) - h_l^*(t, x)| \leqslant \chi(t, x) \quad ([t, x) \in \mathscr{L}, x \in X) \tag{30.8}$$

$$|h_r(x, u) - h_r^*(x, u)| \leqslant \chi(x, u) \quad ([x, u) \in \mathscr{R}, x \in X) \tag{30.9}$$

If the variation on the left of (30.1) is finite, there are $\mathbf{A}_1$, χ_1, satisfying

$$\chi_1(E) < V(\mathbf{h}; E; X_1) + \epsilon \tag{30.10}$$

$$|h_l(t, x)| \leqslant \chi_1(t, x) \quad ([t, x) \in \mathscr{L}_1, x \in X_1) \tag{30.11}$$

$$|h_r(x, u)| \leqslant \chi_1(x, u) \quad ([x, u) \in \mathscr{R}_1, x \in X_1) \tag{30.12}$$

From (30.8; 30.9; 30.11; 30.12),

$$|h_l^*(t, x)| \leqslant \chi(t, x) + \chi_1(t, x) \quad ([t, x) \in \mathscr{L} \cap \mathscr{L}_1, x \in X_1) \tag{30.13}$$

$$|h_r^*(x, u)| \leqslant \chi(x, u) + \chi_1(x, u) \quad ([x, u) \in \mathscr{R} \cap \mathscr{R}_1, x \in X_1) \tag{30.14}$$

From (30.7; 30.10; 30.13; 30.14),

$$V(\mathbf{h}^*; E; X_1) \leqslant \chi(E) + \chi_1(E) < V(\mathbf{h}; E; X_1) + 2\epsilon$$

---

† Result (30.5) is like the Chebyshev inequality for integrals.

t

As $\epsilon > 0$ is arbitrary we obtain

$$V(\mathbf{h}; E; X_1) \geqslant V(\mathbf{h}^*; E; X_1) \qquad (30.15)$$

This is also true when the left-hand side is infinite. Interchanging **h** and **h*** in (30.15) we obtain (30.1). Similarly for (30.2), in which case the intervals in (30.11; 30.12; 30.13; 30.14) lie in an inner cover contained in

$$(\mathscr{L} \cap \mathscr{L}_1) \cup (\mathscr{R} \cap \mathscr{R}_1)$$

Then (30.3) follows from (30.1), since in X_2,

$$|f(x)h_s^*| \leqslant C|h_s^*|\,(s = l, r)$$

and similarly for (30.5). Likewise (30.4; 30.6) follow from (30.2). Name due to Vitali "Sull' integrazione per serie" Rend. del Circ. math. di Palermo (1907), "Sui gruppi di punti e sulle funzioni di variabili reali (Rend. dell' R. Acc. delle Sc. di Torino, 1908)

A pair **h** of interval functions is *absolutely continuous* (*AC**) *with respect to a pair* **h*** *in a set X, relative to E*, if given $\epsilon > 0$, there is a $\delta > 0$ such that for every set X_4 in X with

$$V(\mathbf{h}^*; E; X_4) \leqslant \delta \text{ we have } V(\mathbf{h}; E; X_4) \leqslant \epsilon$$

The pair **h** *is generalized absolutely continuous* (*ACG**) *with respect to* **h*** *in X, relative to E*, if X is a countable union of sets in which **h** is AC^* with respect to **h***, relative to E. If $X = E$ we replace '*in E, relative to E*' simply by '*in E*'.

THEOREM 30.2. *If* **h** *is* AC* *with respect to* **h*** *in* X, *relative to* E, *where* **h*** *is continuous in* E, *and* VB* *in* X, *relative to* E, *then* **h** *is* VB* *in* X, *relative to* E.

p.53 By Theorem 28.5 (28.12), and the similar result for $[x, u)$, the finite variation of **h*** in X, relative to I, tends to 0 with the length of I, for $I = [t, x]$ and $[x, u]$ (*fixed x*); and so for $I = [t, u](t < x < u$, *x fixed*), by Theorem 28.2 (28.2). Then p.50 as in Ex.28.2, p.55 if I is an interval,

$$V(\mathbf{h}^*; E; X \cap I) = V(\mathbf{h}^*; E \cap I; X) \qquad (30.16)$$

so that the former also tends to 0 with the length of I. Dividing E up into a finite number of intervals I for which the right-hand side of (30.16) is δ, which is possible by the continuity, we obtain by absolute continuity,

$$V(\mathbf{h}; E; X \cap I) \leqslant \epsilon$$

which with Theorem 28.1 gives the result. p.49

In Theorems 30.3, 30.4 we assume the existence of

$$H(u, v) = (V)\int_u^v \mathbf{h}, \text{ or } K(u, v) = (V)\int_u^v f\,d\mathbf{h}^*$$

for all intervals $[u, v]$ in an elementary set E.

NV THEOREM 30.3. *The integral* H *is* AC* *with respect to* **h** *in* E. *The integral* K *is* ACG* *with respect to* **h*** *in* E.
We replace **h*** in Theorem 30.1 (30.1) by H to obtain the first p.57 part. For the second we divide E up into sets X_{5j} of x where

$$j-1 \leqslant |f(x)| < j,\ x \in E\,(j = 1, 2, \ldots)$$

and Theorem 30.1 (30.3), with **h** replaced by K, gives what we need.

NV THEOREM 30.4. *If* H *is* VB* (*or* VBG*) *in a set* X, *relative to* E, *then so is* **h**, *and conversely. If* K *is* VBG* *in* X, *relative to* E, *then* **h*** *is* VBG* *in* $X \cap X_6$, *relative to* E, *where* X_6 *is the set where* $f \neq 0$; *and conversely.*
If the pair of interval functions were continuous we could use Theorems 30.2, 30.3. However, it is easy to prove the result directly. For the first part we replace **h*** in Theorem 30.1 (30.1) by H. For the second part, first let K be VBG^* in X, relative to E. Then by definition K is VB^* in each set of a sequence $\{X_{7k}\}$ of mutually disjoint sets with union X. Replacing **h** in Theorem 30.1 (30.5) by K, and X_3 by $X_{7k} \cap X_{8j}$, where X_{8j} is the set where

$$(j-1)^{-1} > |f(x)| \geqslant j^{-1}\,(j = 1, 2, \ldots)$$

we see that **h*** is VB^* in $X_{7k} \cap X_{8j}$, and so VBG^* in $X \cap X_6$. Conversely, let **h*** be VBG^* in $X \cap X_6$. Then **h*** is VB^* in each set of a sequence $\{X_{9k}\}$ of mutually disjoint sets with union $X \cap X_6$. Replacing in Theorem 30.1 (30.3), **h** by K, and X_2 by $X \cap CX_6$, $X_{8j} \cap X_{9k} (j, k = 1, 2, \ldots)$ we see that K is VB^* in each of these sets, and so is VBG^* in X, relative to E.

⑤

31. Special results

THEOREM 31.1. *If* |h| *is a finitely subadditive interval function in* [a, b], *and in particular, if* h *is finitely additive there, then*

$$|\mathrm{h}(a, b)| \leqslant V(\mathrm{h}; [a, b]) \qquad (31.1)$$

Ex 5.9

If |h| *is finitely superadditive then* h *is* VB* *and*

$$|h(a, b)| \geqslant V(h; [a, b]) \tag{31.2}$$

If |h| *is finitely additive then* h *is* VB* *and*

$$|h(a, b)| = V(h; [a, b]) \tag{31.3}$$

In particular, if **h** *is a pair of interval functions,*

$$V(\mathbf{h}; [a, b]) = V(V(\mathbf{h}; .); [a, b]) \tag{31.4}$$

If h *is real, finitely additive and* VB* *in* [a, b], *then* h *is the difference of two* VB* *non-negative and finitely additive interval functions,*

$$h(I) = \tfrac{1}{2}\{V(h; I)+h(I)\}-\tfrac{1}{2}\{V(h; I)-h(I)\} \tag{31.5}$$

If h is not VB^* in (31.1) there is nothing to prove. Thus we can assume that h is VB^*. If h is also finitely additive, then

$$|h(u, w)| = |h(u, v)+h(v, w)| \leqslant |h(u, v)|+|h(v, w)|$$

for $a \leqslant u < v < w \leqslant b$, and $|h|$ is finitely subadditive. If $|h|$ is finitely subadditive, χ finitely superadditive, and

$$|h| \leqslant \chi \tag{31.6}$$

for the intervals of $\mathscr{L} \cap \mathscr{R}$, where A is complete in $[a, b]$, then $\chi-|h|$ is finitely superadditive, and from (21.6), p. 30

$$|h(a, b)| \leqslant \chi(a, b)$$

Taking the infimum of such $\chi(a, b)$, we obtain (31.1).

If $|h|$ is finitely superadditive it is a χ satisfying (31.6), so that h is VB^* and the inequality in (31.2) follows. From this and (31.1) we obtain (31.3).

By Theorem 28.2, p.50, the variation is finitely additive, so that (31.3) gives (31.4), and the interval functions in the curly brackets of (31.5) are finitely additive. These interval functions are non-negative by (31.1).

→

THEOREM 31.2. *If* $|\mathbf{h}| = \{|h_l|, |h_r|\}$ *is variationally integrable in an elementary set* E *then*

√

$$(V)\int_E |\mathbf{h}| = V(\mathbf{h}; E) \tag{31.7}$$

h *is* VB*, *and* V(**h**; .) *is variationally equivalent to* |**h**| *in* E. *If*

$$\{ch(X)|h_l|, ch(X)|h_r|\} \tag{31.8}$$

is variationally integrable in E *then*

$$(V)\int_E ch(X)\, d|\mathbf{h}| = V(\mathbf{h}; E; X) \tag{31.9}$$

which is finite, and V(**h**; .; X) *is variationally equivalent to* (31.8) *in* E.

In Theorem 31.1 (31.1; 31.2; 31.3) h *being* VB* *in* (31.1), *the variation is also the variational integral of* |h| (31.10)

If **h** *is variationally integrable in* E *then*

Longer modulus bars.

$$\left|(V)\int_E \mathbf{h}\right| \leqslant V(\mathbf{h}; E) \tag{31.11}$$

If this variation is finite then |**h**| *is also variationally integrable, and* (31.11) *becomes*

$$\left|(V)\int_E \mathbf{h}\right| \leqslant (V)\int_E |\mathbf{h}| \tag{31.12}$$

True also for complex. See p. 45

The decomposition (31.5) *of a real finitely additive* VB* *interval function* h *is the most economical, in the sense that if* $h = h_1-h_2$, *where* h_j *is* VB*, *non-negative, and finitely additive, for* $j = 1, 2$, *then*

$$h_1 \geqslant \tfrac{1}{2}\{V(h; .)+h\},\ h_2 \geqslant \tfrac{1}{2}\{V(h; .)-h\} \tag{31.13}$$

For (31.7) we use Ex. 26.1, p.47, and Theorems 30.1 (30.1) (p. 57), 21.1 (p. 31) and 31.1 (31.3). Thus

$$V(\mathbf{h}; E) = V(|\mathbf{h}|; E) = V\left((V)\int_{\emptyset} |\mathbf{h}|; E\right) = (V)\int_E |\mathbf{h}|$$

giving ~~all~~ the ~~required~~ results. Then (31.9) ~~is an easy deduction,~~ follows on substituting (31.8) for |**h**| in (31.7), and using Ex. 26.2, p.47. For (31.10) we use Theorem 22.1, p.34, to show the existence of the variational integral of $|h|$ when $|h|$ is either finitely subadditive or finitely superadditive. Then (31.7) gives the result. For (31.11) we use Theorem 21.1, p.31, for the finite additivity of the integral, and then Theorems 31.1 (31.1) and p.57, 30.1 (30.1), obtaining

$$|(V)\int_E \mathbf{h}| \leqslant V\Big((V)\int \mathbf{h}; E\Big) = V(\mathbf{h}; E)$$

If the last variation is finite, then by (26.1; 26.2) , p.46,

$$(\mathscr{D})\sum |h_s|$$

is bounded for all divisions $\mathscr{D}$ of E from some complete set. We now use Theorem 25.2 , p.45, and result (31.7) to obtain (31.12). For (31.13) by taking h positive and then negative, we see that

$$h_1 \geqslant \tfrac{1}{2}(|h|+h),\ h_2 \geqslant \tfrac{1}{2}(|h|-h) \tag{31.14}$$

By (31.10) the integral of $|h|$ exists and is the variation, so that if we integrate the four sides in (31.14) we obtain (31.13), as required.

NJ

Theorem 31.3. *Let* X, X$_1$ *be the sets where* f ≠ 0, f ≠ φ, *respectively, where* f, φ *are point functions in an elementary set* E. *Then*

Th.2.6.1

$$K(u, v) \equiv (V)\int_u^v f d\mathbf{h}^* = 0$$

for all intervals [u, v] *in* E *if and only if*

$$V(\mathbf{h}^*; E; X) = 0 \tag{31.15}$$

If, for all [u, v] *in* E,

$$(V)\int_u^v f\, d\mathbf{h}^* = (V)\int_u^v \varphi\, d\mathbf{h}^* \tag{31.16}$$

then

$$V(\mathbf{h}^*; E; X_1) = 0 \tag{31.17}$$

Conversely, if (31.17) *is true, and if either side of* (31.16) *exists, then both sides exist and are equal. If for all* [u, v] *in* E,

$$(V)\int_u^v f\, d\mathbf{h}^* = (V)\int_u^v f\, d\mathbf{h} \tag{31.18}$$

then

$$V(\{h_1^* - h_1, h_r^* - h_r\}; E; X) = 0 \tag{31.19}$$

Conversely, if (31.19) *is true, and if either side of* (31.18) *exists, then both exist and are equal. If* $\mathbf{h}^*$ *is variationally equivalent to* {φh$_1$, φh$_r$} *in* E, *then*

$$(V)\int_u^v f\, d\mathbf{h}^* = (V)\int_u^v f\varphi\, d\mathbf{h} \tag{31.20}$$

for all intervals [u, v] *in* E, *in the sense that if one side of* (31.20) *exists then so does the other, and they are equal. If* {fh$_1$, fh$_r$} *and* |**h**| *are variationally integrable, the former being* VB*, *in* E, *then*

$$|(V)\int_E f\, d\mathbf{h}| \leqslant (V)\int_E |f|\, dV(\mathbf{h}; .) \tag{31.21}$$

This is a most important theorem. In particular it enables us to relax the definitions of the variational and Riemann-complete integrals, in that we can replace the point function f in K by any other point function φ equal to f except on a set in which the variation of **h***, relative to E, is zero. If **h*** is understood, we can call the set a set of *variation zero*, and we can say that $f = \varphi$ *almost everywhere* (abbreviated to *a.e.*).

Result (31.20) corresponds to the calculus process of *integration by substitution*, Section 2 (2.8), p. 5). Finally, inequality (31.21) is well known for the special case

$$|\int_E f\,dx| \leqslant \int_E |f|\,dx$$

The inequality (31.21) follows from Theorem 31.2 (31.12; 31.7) and result (31.20). In turn, if **h*** is variationally equivalent to $\{\varphi h_l, \varphi h_r\}$ then (31.19) is true with **h** replaced by $\{\varphi h_l, \varphi h_r\}$; and (31.18) is then (31.20). Thus (31.20) follows from the implication of (31.18) by (31.19). Further, by Theorem 19.1, p. 27, the variational integral is a distributive functional. It follows that all the results now hinge on the first result and its converse, to the proof of which we turn.

We first assume (31.15). Excluding $f = 0$, let X_{2j} be the set where

$$j-1 < |f(x)| \leqslant j\ (j = 1, 2, \ldots)$$

p.49, By Theorem 28.1, K for **h** in Theorem 30.1 (30.3), p.57, and Ex. 26.6, p.48,

$$V(K; E) = V(K; E; E) \leqslant V(K; E; E \cap CX) + \sum_{j=1}^{\infty} V(K; E; X_{2j}) \leqslant \sum_{j=1}^{\infty} jV(\mathbf{h}^*; E; X_{2j}) = 0$$

J, so that there is a point x_1 of X_{2n} such that either $x_1 = a_1$ or

$$V(\mathbf{h}; [a, b]; X \cap [a_1, x_1]) < \eta \tag{32.7}$$

Let $[x_1, u_1]$ be in $\mathscr{F}$ with $u_1 - x_1 > 1/n$. If points of X_{2n} lie to the right of u_1, let a_2 be their infimum. Then either $a_2 = x_2$ in X_{2n}, or there is a point x_2 of X_{2n} such that

$$V(\mathbf{h}; [a, b]; X \cap [a_2, x_2]) < \eta \tag{32.8}$$

Let $[x_2, u_2]$ be in $\mathscr{F}$ with $u_2 - x_2 > 1/n$, and so on. If in N such steps we are at b_1, then by (32.5),

$$(N-1)/n < l,\ N < nl+1,\ N\eta < \tfrac{1}{2}\epsilon \tag{32.9}$$

Using Theorem 28.1 and (32.4; 32.7; 32.8; ...), and putting

$$X_1 = \bigcup_{j=1}^{N} [x_j, u_j],\ X_3 = \bigcup_{j=1}^{N} [a_j, x_j], \text{ then } X_{2n} \subseteq X_1 \cup X_3$$

$$\begin{aligned} V(\mathbf{h}; [a, b]; X) - \tfrac{1}{2}\epsilon &< V(\mathbf{h}; [a, b]; X_{2n}) \\ &\leqslant V(\mathbf{h}; [a, b]; X \cap X_1) + \sum_{j=1}^{N} V(\mathbf{h}; [a, b]; [a_j, x_j]) \\ &\leqslant V(\mathbf{h}; [a, b]; X \cap X_1) + N\eta \end{aligned}$$

Then (32.9) gives (32.1). For (32.2), we use Theorems 28.2 (28.2) and 31.2 (31.7), obtaining

50, p.60,

$$V(\mathbf{h}; X_1) = \sum_{j=1}^{N} V(\mathbf{h}; [x_j, u_j])$$

$$= \sum_{j=1}^{N} L(x_j, u_j) \leqslant \sum_{j=1}^{N} \chi(x_j, u_j) \leqslant \chi(a, b)$$

In particular, we can take the χ to correspond to the inner variation, splitting up the set X into sets X_4, X_5 in which, respectively (27.1) is true and (27.2) is true. Let X_6 and X_7 be the corresponding sets X_1. Then by (32.1), continuity and (32.2),

$$\begin{aligned} V(\mathbf{h}; E; X) &\leqslant V(\mathbf{h}; E; X_4) + V(\mathbf{h}; E; X_5) \\ &< V(\mathbf{h}; E; X_4 \cap X_6) + V(\mathbf{h}; E; X_5 \cap X_7) + 2\epsilon \\ &\leqslant V(\mathbf{h}; E; X_6) + V(\mathbf{h}; E; X_7) + 2\epsilon \\ &= V(\mathbf{h}; X_6) + V(\mathbf{h}; X_7) + 2\epsilon \\ &\leqslant 2\,.\,IV(\mathbf{h}; E; X) + 4\epsilon \end{aligned}$$

As $\epsilon > 0$ is arbitrary, we have (32.3).

Next comes the question of how much will remain true if $\mathbf{h}$ is no longer continuous. As a preliminary we need a theorem that is valid also in abstract spaces.

Theorem 32.2. *If* $\mathbf{h}$ *is* VB* *and variationally integrable in an elementary set* E *then* $\mathbf{h} = \{\mathrm{h}_l{}^* + \mathrm{h}_e, \mathrm{h}_r{}^* + \mathrm{h}_e\}$, *where* $\mathbf{h}^*$ *is continuous in* E *and* VB* *and variationally integrable there, and* h_e *is finitely additive with discontinuities at a set of points in* E *that is at most countable, say, as the sequence* $\{\mathrm{j}_\mathrm{n}\}$. *We can write*

$$\mathrm{h}_e\,(\mathrm{a}, \mathrm{b}) = \Sigma'\, \mathrm{a}_\mathrm{n} + \Sigma''\, \mathrm{b}_\mathrm{n}$$

where Σ', Σ'' *are respectively over those* n *for which*

$$\mathrm{a} < \mathrm{j}_\mathrm{n} \leqslant \mathrm{b},\ \mathrm{a} \leqslant \mathrm{j}_\mathrm{n} < \mathrm{b},$$

the two series being absolutely convergent.

Let H be the variational integral of $\mathbf{h}$. By Theorem 21.2 (21.12; 21.13), p. 33, $\{h_l - H, h_r - H\}$ is continuous in E, so that we can assume that $\mathbf{h} = H$, finitely additive. Further, by Theorem 31.1 (31.5), p. 60, we can assume that H is non-negative, so that $H(a, t)$ is monotone increasing in t, for $[a, t]$ in E, and is bounded above by $H(E)$. Using Theorem 13.2, p.17, $H(a, t)$ tends to a limit as $t \to x-$, for each fixed x such that $[a, x]$ is in E. Then by finite additivity we see that, for such an x,

$$H(x-) = \lim_{t \to x-} H(t, x)$$

exists. Similarly, if $[x, b]$ is in E for some $b > x$,

$$H(x+) = \lim_{u \to x+} H(x, u)$$

exists. If now we take n non-overlapping intervals $[t_k, x_k]$ in E we obtain

$$\sum_{k=1}^{n} H(x_k-) \leqslant \sum_{k=1}^{n} H(t_k, x_k) \leqslant H(E) \tag{32.10}$$

by finite additivity and non-negativeness. It follows that the number of points x_k for which $H(x_k-) \geqslant 1/m$, cannot be greater than $mH(E)$, and similarly for $H(x_k+)$. Taking $m = 1, 2, \ldots$, we obtain an at most countable set of points j_n

By Theorem 21.1, p. 31, K is finitely additive. Then Ex. 26.6 p.48 and Theorem 31.1 (31.1) give $K(u, v) = 0$ for all $[u, v]$ in E. Conversely, if $K(u, v) = 0$ for all $[u, v]$ in E then by Theorem 30.1 (30.5), p.57, using the X_{8j} of Theorem 30.4, p.59,

$$V(\mathbf{h}^*; E; X_{8j}) \leqslant jV(K; E; X_{8j}) = 0$$

Theorem 28.1 p.49 completes the proof.

Ex. 31.1. If $h_s^* \geqslant 0 (s = l, r), f \geqslant 0, K(E) = 0$, then (31.15) is true.

(For K is finitely additive and non-negative, so that $K(u, v) = 0$ for all $[u, v] \subseteq E$.)

Ex. 31.2. If

Ex 15.3

$$h(I) = \begin{cases} mI \ (mI \text{ rational}) \\ 0 \ (mI \text{ irrational}) \end{cases}$$

where mI is the length of the interval I, show that

$$V(h; I) = mI, \quad IV(h; I) = 0$$

and that h is not variationally equivalent to $V(h; I)$ in [0, 1].

Ex 31.3. If f is bounded and variationally integrable with respect to $\mathbf{h}$ in an elementary set E, with $f = 0$ except in a set X, prove that

$$\left|(V) \int_E f \, d\mathbf{h}\right| \leqslant \sup_X |f| \, . \, V(\mathbf{h}; E; X)$$

(Use Theorem 31.2 (31.11); and Theorem 30.1 (30.3), with $f = fch(X)$.)

Ex. 31.4. Prove (31.21), assuming only that the two integrals exist.

(Use Theorem 28.4 with a $\chi_j(E) < \epsilon.4^{-j}$ for points x with $2^{j-1} \leqslant |f(x)| < 2^j$.) $. \ |\Sigma f h_s| \leqslant \Sigma |f h_s| \leqslant \Sigma \{|f| V(h; \bullet; X_j) + |f| \chi_j\} \leqslant \Sigma |f| V(h; \cdot) + \Sigma 2^j \chi_j(E)$. Last $\Sigma < \epsilon$.

32. Sierpiński's lemma and the inner variation

Previous theorems given in the book, except Theorems 22.2, p.34, and 23.1, p.35 can be generalized to hold in more than one dimension, and sometimes in an abstract space. This section gives results that are essentially one-dimensional and that are not valid for Euclidean spaces with finite dimension greater than 1. Theorem 32.1 was first given by W. Sierpinski (1923) using outer Lebesgue measure.

Theorem 32.1. *Let $\mathscr{F}$ be a family of closed intervals contained in an elementary set E, and let X be a set in E, such that to each point x of X there corresponds an interval [x, u] of $\mathscr{F}$. Let **h** be continuous, and VB* in X, relative to E. Then given $\epsilon > 0$, there is a finite number of disjoint intervals of $\mathscr{F}$ with union X_1, such that*

$$V(\mathbf{h}; E; X \cap X_1) > V(\mathbf{h}; E; X) - \epsilon \tag{32.1}$$

*Similarly for left-hand intervals. If |**h**| is variationally integrable in E with integral L, if χ is non-negative and finitely superadditive, and if $L \leqslant \chi$ for intervals of $\mathscr{F}$, then*

$$V(\mathbf{h}; X_1) \leqslant \chi(E) \tag{32.2}$$

unnecessary, by Th. 31.2 p.60

*If **h** is continuous ~~and VB* in X, relative to E~~, with |**h**| variationally integrable in E, then*

$$IV(\mathbf{h}; E; X) \leqslant V(\mathbf{h}; E; X) \leqslant 2.IV(\mathbf{h}; E; X) \tag{32.3}$$

so that in particular, sets of inner variation zero are also of variation zero.

Note that Ex. 31.2 shows that if |**h**| is not variationally integrable, the last result may be false.

We can obviously assume that E is an interval $[a, b]$. Then the general case follows by finite union. Let X_{2n} be the set of x in X for which there is an $[x, u]$ of $\mathscr{F}$ with $u - x > 1/n$. Then

$$X = \lim_{n \to \infty} X_{2n}$$

so that by Theorem 28.6, p. 54, we can choose n such that

$$V(\mathbf{h}; [a, b]; X_{2n}) > V(\mathbf{h}; [a, b]; X) - \tfrac{1}{2}\epsilon \tag{32.4}$$

We now put

$$a_1 = \inf X_{2n} \geqslant a, \quad b_1 = \sup X_{2n} \leqslant b, \quad l = b_1 - a_1, \qquad \eta = \tfrac{1}{2}\epsilon/(nl+1) \tag{32.5}$$

By Theorem 28.5, p.53 $V(\mathbf{h}; I)$ is continuous in I; and if the end-points of an interval J are inside I then

$$V(\mathbf{h}; [a, b]; X \cap J) \leqslant V(\mathbf{h}; I) \tag{32.6}$$

It follows that the first variation in (32.6) is continuous in at which either $H(j_n-) > 0$ or $H(j_n+) > 0$, or both. At all other points, $H(x-) = H(x+) = 0$; and from (32.10),

$$\sum_{n=1}^{\infty} \{H(j_n-)+H(j_n+)\} \leqslant 2H(E)$$

Actually, the 2 can be removed, but what has been proved is sufficient to give the h_e with the required properties, if we put $a_n = H(j_n-)$, $b_n = H(j_n+)$ when **h** is non-negative.

Theorem 32.3. *Let* **h** *be* VB* *and variationally integrable in* E, *but not necessarily continuous. Then Theorem* 32.1 (32.1) *is still true, while* (32.3) *can be false. However,* (32.3) *is true if* $|\mathrm{H}(\mathrm{x}-)| = |\mathrm{H}(\mathrm{x}+)|$ *at every point; while a zero inner variation implies a zero variation if* $\mathrm{H}(\mathrm{x}-)$, $\mathrm{H}(\mathrm{x}+)$ *are only zero together.*

For the first part we take the integer M so great that

$$\sum_{n=M}^{\infty} \{|H(j_n-)|+|H(j_n+)|\} < \tfrac{1}{2}\epsilon$$

where H is the variational integral, $\{j_n\}$ is the sequence of discontinuities of H, and $H(j_n\pm)$ are as in Theorem 32.2. To each j_n in X, with $n < M$, there corresponds an interval $[j_n, u_n]$ of $\mathscr{F}$. From these we can take a disjoint collection $\mathscr{F}_1$ such that each j_n in X with $n < M$ lies in an interval of the collection, and then we remove the interiors of those intervals from E to obtain a new elementary set for which we prove Theorem 32.1 using $\mathbf{h}^*$ for **h** and $\frac{1}{2}\epsilon$ for $\frac{1}{2}\epsilon$. The resulting X_1, together with the intervals of $\mathscr{F}_1$, gives the result for **h**.

To show that (32.3) may be false, it is sufficient to take the case of

$$h(a, b) = \begin{cases} 1 \text{ if } a \leqslant 0 < b \\ 0 \text{ otherwise} \end{cases}$$

If X and E are $[-1, +1]$ then the variation is 1, while the inner variation is 0, the latter being obtained by using a left-hand interval only at 0.

If the inner variation is 0, and if $H(x-)$, $H(x+)$ can only be zero together, then no discontinuity j_n can lie in X, so that by Theorem 32.1,

$$\begin{aligned} 0 = IV(\mathbf{h}; E; X) &= IV(\mathbf{h}^*; E; X) \\ &= V(\mathbf{h}^*; E; X) = V(\mathbf{h}; E; X) \end{aligned}$$

(McGrotty) FALSE. Take $h(u,v) = \begin{cases} 1 & (u < 0 \leq v) \\ 0 & (\text{otherwise}) \end{cases}$, $E = X = [-1, 1]$, each point of E has a right-hand interval $[x, v]$, with $v < 0$ if $x < 0$, and $h(x,v) = 0$. $V(h; E; X_1) = 0$, $V(h; E) = 1$.

If the inner variation is not 0, and if $|H(x-)| = |H(x+)|$, we can enclose the discontinuity j_n in an interval in which the variation of $\mathbf{h}^*$ is less than 2^{-m-n}, and letting $m \to \infty$, and using Theorem 32.1,

$$\begin{aligned} V(\mathbf{h}; E; X) &= V(\mathbf{h}^*; E; X) + \textstyle\sum_1\{|H(j_n-)|+|H(j_n+)|\} \\ &\leqslant 2IV(\mathbf{h}^*; E; X) + \textstyle\sum_1 2|H(j_n-)| = 2IV(\mathbf{h}; E; X) \end{aligned}$$

where Σ_1 is the sum over all n with j_n in X.

Ex. 32.1. If $|\mathbf{h}|$ is variationally integrable in E, show that when the intervals of $\mathscr{F}$ can be taken from a given set complete in E, we can also assume that

$$V(\mathbf{h}; X_1) < V(\mathbf{h}; E; X)+\epsilon$$

(Take χ to correspond to the variation of **h**.)

33. Integration by parts

Here we consider properties, essentially one-dimensional, that generalize (2.7) (p. 4). If f is a function of points we put

$$H_f(u, v) = f(v)-f(u)$$

Theorem 33.1. *If the integrals in*

VV

$$(\mathrm{V}) \int_{\mathrm{u}}^{\mathrm{v}} \mathrm{f}\, d\mathrm{g} + (\mathrm{V}) \int_{\mathrm{u}}^{\mathrm{v}} \mathrm{g}\, d\mathrm{f} = \mathrm{H}_{\mathrm{fg}}(\mathrm{u}, \mathrm{v}) \tag{33.1}$$

exist, and if the equality is true, for each [u, v] *in an elementary set* E, *then*

$$\mathrm{V}(\mathrm{H_f H_g}; \mathrm{E}) = 0 \tag{33.2}$$

Conversely, if (33.2) *is true, and if the first integral in* (33.1) *exists, so does the second, and we have equality in* (33.1).

p.57, Theorem 30.1 (30.1), and the identity

$$\begin{aligned} f(x)\{g(x)-g(t)\}+g(x)\{f(x)-f(t)\}-\{f(x)g(x)-f(t)g(t)\} \\ =\{f(x)-f(t)\}\{g(x)-g(t)\} \end{aligned} \tag{33.3}$$

show that if the integrals in (33.1) exist, and if (33.1) is true,

$$\begin{aligned} V(H_f H_g; E) &= V(fH_g+gH_f-H_{fg}; E) \\ &= V\Big((V)\int f\,\mathrm{d}g+(V)\int \mathrm{g}\,\mathrm{d}f-H_{fg}; E\Big) = 0 \end{aligned}$$

Conversely, if the first integral in (33.1) exists, then from (33.2; 33.3) and Theorem 30.1 (30.1), p.57,

$$V\left(gH_f+\left\{(V)\int f\,dg-H_{fg}\right\};E\right)=0$$

which proves the existence of the second integral in (33.1) and the equality.

Ex 15.7 p.147

THEOREM 33.2. *If* (33.2) *is true then* f *is continuous except in a set* X *with* If V(HfHg; E)=0 then

$$IV(H_g;E;X)=0 \tag{33.4}$$

If g *is continuous, or discontinuous on both sides, at each point of* E, *and* VB* *in* X, *relative to* E, *the variation is also* 0. *But neither integral in* (33.1) *need exist.*

If for fixed $\epsilon>0$, $X_1(\epsilon)$ is the set of all points x of discontinuity of f with the property that there is a sequence of t tending to x such that

$$|f(t)-f(x)|\geqslant\epsilon$$

then for $[t,x)$ (when $t<x$), or $[x,t)$ (when $t>x$), we have

$$\epsilon|H_g|\leqslant|H_fH_g|,\ IV(H_g;E;X_1(\epsilon))=0$$

Taking $\epsilon=1/n$, $n=1,2,\ldots$, the union of the $X_1(\epsilon)$ is clearly X, and by Theorem 28.1 p.49, it satisfies (33.4). If g is continuous, or discontinuous on both sides, at each point of E, and VB^* in X, relative to E, then Theorem 32.1 (32.3) and 32.3 show that the variation is also 0. To prove the third part we take $g(x)=x$ (*all* x) and

$$f(x)=-1/(x\log x)\ \ (0<x\leqslant\tfrac{1}{2}),\ \ f(0)=0$$

By Ex. 33.1 at the end of this section, for each ϵ in $0<\epsilon<\frac{1}{2}$,

$$V(H_fH_g;[\epsilon,\tfrac{1}{2}])=0,\ \ V(H_fH_g;[0,\tfrac{1}{2}])=V(H_fH_g;[0,\epsilon]) .$$

By Theorem 28.5 (28.12), p.53

$$\begin{aligned}V(H_fH_g;[0,\tfrac{1}{2}])&=\lim_{\epsilon\to0}V(H_fH_g;[0,\epsilon])\\&=\limsup_{\epsilon\to0}|H_f(0,\epsilon)H_g(0,\epsilon)|\\&=\limsup_{\epsilon\to0}(-1/\log\epsilon)=0\end{aligned}$$

and (33.2) is satisfied with $E=[0,\frac{1}{2}]$. On the other hand, for $x=e^{-u}$,

$$(NL)\int_\epsilon^{\frac{1}{2}}f(x)\,dx=(NL)\int_{\log 2}^{\log(1/\epsilon)}\frac{du}{u}=\log\log(1/\epsilon)-\log\log 2$$

and Ex. 17.2 (p. 25), shows that the Riemann-complete integral has the same value. It follows that the Riemann-complete integral in $[\epsilon,\frac{1}{2}]$ does not tend to a limit as $\epsilon\to0$, so that by Theorem 21.2 p.32, the Riemann-complete integral cannot exist in $[0,\frac{1}{2}]$. Thus, using Theorem 33.1, neither integral

$$(V)\int_0^{\frac{1}{2}}f\,dx,\quad (V)\int_0^{\frac{1}{2}}x\,df$$

can exist, proving the third part of the theorem.

If X is a set of points we denote by X', *the derived set of* X, that is, the set of all limits of convergent sequences of distinct points of X. We also say that a point function f *is continuous in a set* X, if for every sequence of points t of X, with limit x, the sequence of $f(t)$ tends to $f(x)$. Further, if f, g are point functions, and if x is the associated point, and t the other end-point, of an interval, we denote by g^*H_f the pair of interval functions

$$g(t)\{f(x)-f(t)\}(t<x),\ g(t)\{f(t)-f(x)\}(t>x)$$

The following theorem has connections with Theorem 22.2 p.34

THEOREM 33.3.

Ex.15.9

$$(V)\int_u^v f\,dg=H_{fg}(u,v) \tag{33.5}$$

for all [u, v] *in an elementary set* E *if and only if*

$$V(g^*H_f;E)=0 \tag{33.6}$$

In particular, if X_{2n} *is the set of* x *in* E *for which* $|g(x)|\geqslant 1/n$, *then*

$$f \textit{ is continuous in } X_{2n}\cup X'_{2n}=\overline{X}_{2n} \tag{33.7}$$

$$IV(H_f;E;X'_{2n})=0 \tag{33.8}$$

If f *is continuous, or discontinuous on both sides, at each point of* E, *and* VB* *in* E, *then*

$$V(H_f;E,X'_{2n})=0 \tag{33.9}$$

If [a, b] $\subseteq X'_{2n}$, *and if* (33.5) *is true for all* [u, v] $\subseteq$ [a, b], *then* f *is constant in* [a, b]. (33.10)

For $t < x$ or $t > x$ we have the identity

$$f(x)\{g(x)-g(t)\}-\{f(x)g(x)-f(t)g(t)\} = g(t)\{f(t)-f(x)\},$$

$$fH_g - H_{fg} = -g^*H_f \tag{33.11}$$

When the variational integral of f with respect to g exists in E and satisfies (33.5), we use (33.11) and Theorem 30.1 (30.1)
p.57 to obtain

$$V(g^*H_f; E) = V\left(H_{fg} - (V)\int f\,\mathrm{d}g; E\right) = 0$$

Conversely, from (33.6; 33.11), fH_g is variationally equivalent to the finitely additive H_{fg}, proving (33.5) with the existence of the integral.

From (33.6),

$$\lim_{t\to x} g(t)\{f(t)-f(x)\} = 0 \;\; (t, x \textit{ in } E)$$

Taking t in X_{2n}, $\underline{x}$ in X'_{2n}, we see that f is continuous in X_{2n}. Further, if t_j in $\bar{X}_{2n}$ tends to a point x, there is a sequence of points u_j in X_{2n} such that

$$|u_j - t_j| < 1/j, \; |f(u_j) - f(t_j)| < 1/j$$

so that u_j also tends to x, $f(u_j)$ tends to $f(x)$, and so $f(t_j)$ tends to $f(x)$. This gives (33.7). The choice of t in X_{2n}, x in X'_{2n}, also gives (33.8) from (33.6). Then (33.9) follows from (33.8) by Theorems 32.1 (32.3) and 32.3. In the conditions of (33.10), f is continuous in $[a, b]$ by (33.7). By (33.8), taking $E = [a, b]$, we have $IV(H_f; [a, b]) = 0$. More exactly, given $\epsilon > 0$, there is a non-negative finitely superadditive interval function χ such that $\chi(a, b) < \epsilon$, and such that
p.48 (27.1; 27.2) are true with $h_l = h_r = H_f$, with x in $[a, b]$, and with t, u in X_{2n}, everywhere dense in $[a, b]$. By Theorem 16.2 p.23 we obtain

$$|H_f(a, b)| \leqslant \chi(a, b) < \epsilon$$

Hence $f(a) = f(b)$, so that by taking sub-intervals of $[a, b]$ we prove (33.10).

This theorem is nearly a blend of Theorems 33.1 and 31.3 p.62 It seems difficult to break down (33.6) into conditions like (33.7; 33.8) that are sufficient for (33.6) as well as being necessary.

Ex. 33.1. If the derivatives f', g' of f, g exist and are finite in an elementary set E, show that (33.2) is satisfied. Thus Theorem 33.1 covers (2.7) (p. 4), the integration by parts of the calculus.

$$\left(|H_f H_g| \leqslant (|f'|+1)(|g'|+1)(t-x)^2 < \epsilon|t-x| \textit{ for } t \textit{ near to } x\right)$$

Ex. 33.2. If f is continuous and g is VB^*, show that (33.2) is true.

Ex. 33.3. If f is monotone increasing and g is continuous, then prove that in each interval $[a, b]$ there is a point ξ such that

$$(V)\int_a^b f\,\mathrm{d}g = f(a)\{g(\xi)-g(a)\}+f(b)\{g(b)-g(\xi)\}$$

(First prove that the integral exists. Then show that (33.2) is true, integrate by parts, and use Ex.20.4 (p. 29).)

Ex. 33.4. If g is 0 at the rationals and 1 at the irrationals, and if (33.5) exists, then f is constant.

4

DIFFERENTIATION

34. Differentiation of VB* and VBG* functions

The inverse operation to integration is simpler to handle, so that we can obtain more detailed results. Let x be a point in an elementary set E, and let $\mathbf{h}, \mathbf{h}^*$ be two pairs of interval functions. If a number D exists with the property that, given $\epsilon > 0$, there is a $\delta > 0$ such that for intervals in E with length less than δ and with associated point x,

$$|h_s - D.h_s^*| \leqslant \epsilon |h_s^*| \; (s = l, r) \tag{34.1}$$

then D is called the *derivative* $D(\mathbf{h}, \mathbf{h}^*; E; x)$ of $\mathbf{h}$ *with respect to* $\mathbf{h}^*$, E *at* x. If x is at an end of an interval of E, so that one side is not in E, then s takes only one value. Otherwise E makes no difference to the definition of D. If different D, D' satisfy (34.1), then given $\epsilon > 0$, there is a $\delta_1 > 0$ such that for intervals in E with length less than δ_1 and with associated point x, we have

$$|D - D'|.|h_s^*| \leqslant 2\epsilon |h_s^*| \, (s = l, r)$$

Taking $\epsilon = |D - D'|/4$, we see that for these intervals,

$$h_s^* = 0 \; (s = l, r) \tag{34.2}$$

Otherwise D is uniquely defined if it exists. The set of points x satisfying (34.2) for the given intervals, with δ_1 varying with x, is a set in which $\mathbf{h}^*$ has variation zero, relative to E, and so may be disregarded in the sequel.

THEOREM 34.1. *If* $\mathbf{h}^*$ *is arbitrary, while* $\mathbf{h}$ *is* VB* *or* VBG* *in a set* X, *relative to an elementary set* E, *then the ratios*

$$h_l(t, x)/h_l^*(t, x),\ h_r(x, u)/h_r^*(x, u) \quad ([t, x) \subseteq E, [x, u) \subseteq E) \tag{34.3}$$

are either of the form 0/0 *or are bounded as* t, u *tend to* x, *except for* x *in a set* X_1 *contained in* X *such that*

$$IV(\mathbf{h}^*; E; X_1) = 0 \tag{34.4}$$

If $V(\mathbf{h}; E; X) = 0$ *then* $D(\mathbf{h}, \mathbf{h}^*; E; x) = 0$ *in* X, *except for a set* X_2 *satisfying* (34.4). (34.5)

p.55

By using Ex. 28.1 for IV and the sets X_1 we need only prove (34.4) when $\mathbf{h}$ is VB^*. Thus given $\epsilon > 0$, there are a set A complete in E and a non-negative finitely superadditive interval function χ such that for intervals of $\mathscr{L} \cup \mathscr{R}$ with associated point x in X,

$$|h_s| \leqslant \chi \tag{34.6}$$

$$\chi(E) < V(\mathbf{h}; E; X) + \epsilon \tag{34.7}$$

If, for some fixed $M > 0$, and the intervals of some inner cover from $\mathbf{A}$, of some set X_3,

$$|h_s| > M|h_s^*| \tag{34.8}$$

then from (34.6; 34.7; 34.8),

$$M|h_s^*| < \chi,$$

$$M.IV^*(\mathbf{h}^*; E; X_3; \mathbf{A}) \leqslant \chi(E) < V(\mathbf{h}; E; X) + \epsilon$$

$$M.IV(\mathbf{h}^*; E; X_3) \leqslant V(\mathbf{h}; E; X) \tag{34.9}$$

as $\epsilon \to 0$ and $\mathbf{A}$ shrinks. As $X_1 \subseteq X_3$ for each $M > 0$ we take M arbitrarily large and prove (34.4). Under the conditions of (34.5) then (34.9) shows that (34.4) is true for X_3 and each $M > 0$. Taking M arbitrarily *small* we prove (34.5).

Note that when $\mathbf{h}^*$ is continuous in E and VB^* in X, relative to E, with $|\mathbf{h}^*|$ variationally integrable in E, then using Theorem 32.1 (32.3), (34.4) gives

p65

$$V(\mathbf{h}^*; E; X_1) = 0 \tag{34.10}$$

This also holds if we relax continuity to the requirement that

$$\lim_{t \to x-} h_l(t, x)$$

and

$$\lim_{u \to x+} h_r(x, u)$$

are only zero together.

THEOREM 34.2. *Let* $\mathbf{h}$ *be variationally equivalent to* $\mathbf{h}_1$, $\mathbf{h}^*$ *to* $\mathbf{h}_1^*$, *in a set* X, *relative to an elementary set* E. *Then either*

$$D(\mathbf{h}, \mathbf{h}^*; E; x) = D(\mathbf{h}_1, \mathbf{h}_1^*; E; x) \tag{34.11}$$

or else both derivatives do not exist; except for x *in* X_4 *satisfying* (34.4) *with* X_1 *replaced by* X_4.

We can substitute $\{h_l-h_{1l}, h_r-h_{1r}\}$ for **h** in Theorem 34.1 (34.5), which gives the result when $\mathbf{h}^* = \mathbf{h}_1^*$. Thus by Ex. 28.1, p. 55, for IV we need only prove the theorem in the case $\mathbf{h} = \mathbf{h}_1$. If

$$|h_s-D.h_s^*| \leqslant \epsilon|h_s^*|$$

then

$$|\mathbf{h}_s-D.h_{1s}^*| \leqslant \epsilon|h_{1s}^*|+(|D|+\epsilon)|h_s^*-h_{1s}^*|$$

l/ By substituting $|h_s^*-h_{1s}^*|$ for $h_s(s = l, r)$ in Theorem 34.1 (34.5), and $\mathbf{h}_1^*$ for $\mathbf{h}^*$, we prove that if $D(\mathbf{h}, \mathbf{h}^*; E; x)$ exists then so does $D(\mathbf{h}, \mathbf{h}_1^*; E; x)$, with the same value, except for x in a set X_5 satisfying $IV(\mathbf{h}_1^*; E; X_5) = 0$, and so (34.4) with X_1 replaced by X_5, by Theorem 30.1 (30.2) (p. 57). Interchanging $\mathbf{h}^*$ and $\mathbf{h}_1^*$ we prove the theorem.

Also needs Theorem 32.2 p. 67, with $D(h_e, h^*; E; x) = 0$ a.e.

THEOREM 34.3. *Let* **h**, |**h**|, **h*** *be variationally integrable,* **h*** *continuous, and* $h_s^* \geqslant 0(s = l, r)$, *all in an elementary set* E. *Then* D(**h**, **h***; E; x) *exists in* E *except for a set* X_8 of x *satisfying* (34.10).

By Theorem 34.2 we can replace **h**, **h*** by their variational integrals, respectively, H, H^*. By Theorem 31.2 (31.7), (p. 60), **h** is VB^*. By Theorem 30.1 (30.1), (p. 57), so is H. Hence by Theorem 31.1 (31.5) (p. 60), we can suppose that $H \geqslant 0$.

We have reduced the theorem to the case when $H \geqslant 0$, $H^* \geqslant 0$, and both are finitely additive. Given a number $M > 0$, suppose that a set X_7 has the property that for arbitrary A_1 complete in E, and some inner cover of X_7 from $\mathscr{R}_1$, we have $H \geqslant MH^*$. We wish to prove that

$$V(H; E; X_7) \geqslant MV(H^*; E; X_7) \tag{34.12}$$

Given $\epsilon > 0$, there are A_2 complete in E and a non-negative finitely superadditive interval function χ_1 such that for intervals of $\mathscr{L}_2 \cup \mathscr{R}_2$, with the associated point in X_7,

$$0 \leqslant H \leqslant \chi_1 \tag{34.13}$$

$$\chi_1(E) < V(H; E; X_7)+\epsilon \tag{34.14}$$

We take A_2 for A_1 and use Theorem 32.1 (32.1; 32.2;) (p. 65) and (34.13; 34.14). Thus there is a finite collection $\mathscr{Q}$ of closed disjoint intervals from the inner cover, with union U, such that

$$\begin{aligned} M\{V(H^*; E; X_7)-\epsilon\} &\leqslant MV(H^*; E; X_7 \cap U) \leqslant MV(H^*; E; U) \\ &= MV(H^*; U) = M(\mathscr{Q})\, \Sigma\, H^* \leqslant (\mathscr{Q})\, \Sigma\, H \\ &\leqslant (\mathscr{Q})\, \Sigma\, \chi_1 \leqslant \chi_1(E) < V(H; E; X_7)+\epsilon \end{aligned}$$

Letting $\epsilon \to 0$ we prove (34.12). A similar result will hold if the inner cover is from $\mathscr{L}_2$ every time. Hence if for X_7 we have $H \geqslant MH^*$ for one inner cover, and $H \leqslant NH^*$ for another inner cover, each from a fixed side of the points, where $M > N$, then by (34.12) we have

$$MV(H^*; E; X_7) \leqslant V(H; E; X_7) \leqslant NV(H^*; E; X_7)$$

which can only be true if

$$V(H^*; E; X_7) = 0 \tag{34.15}$$

Thus each set for which

$$M = (m+1)2^{-n},\ N = m.2^{-n}\ (m, n = 0, 1, 2, \ldots)$$

satisfies (34.15), so that the union does. In the complement the derivative exists. Thus we have proved the theorem.

THEOREM 34.4. *If* **h** *is variationally integrable and* AC* *with respect to* **h*** *in an elementary set* E, where **h*** *is variationally integrable and continuous and* $h_s^* \geqslant 0(s = l, r)$ *in* E, *then* D(**h**, **h***; E, x) *exists in* E *except for a set* ~~X_8~~ *of* x *satisfying* (34.10), *with* X_1 *replaced by* X_8.

By Theorem 31.2 (p. 60), **h*** is VB^*, so that by Theorem 30.2 (p. 58), **h** is also VB^*. As it is variationally integrable, so is |**h**|, by Theorem 31.2 (p. 60) again. Hence we can use Theorem 34.3 and obtain what we want.

THEOREM 34.5. *If in Theorem* 34.4 *we have* $h_l = h_r = h$, *finitely additive, then*

$$h = (V) \int f\, d\mathbf{h}^*$$

where f = 0 *in* X_8 *and otherwise* f = D (**h**, **h***; E, x); *and* |f| *is variationally integrable with respect to* **h*** *in* E.

For by Theorems 28.1 (p. 49) and 34.4 and the definition (34.1) of differentiation,

$$V(h-f\mathbf{h}^*; E) \leqslant V(h-f\mathbf{h}^*; E; X_8)+V(h-f\mathbf{h}^*; E; CX_8)$$
$$= V(h; E; X_8)$$

This last is 0 as h is AC^* with respect to $\mathbf{h}^*$. Hence the finitely additive h is variationally equivalent to $f\mathbf{h}^*$, proving the first part. The second part of the theorem follows from Theorem 31.2 (p. 60).

Note that in Theorem 34.4 we can change AC^* to ACG^* and obtain that $\mathbf{h}$ is VBG^*. If we can prove that $|\mathbf{h}|$ is variationally integrable, then (34.10) and Theorem 34.5 follow, giving a converse to Theorem 30.3 (p. 59), second part for the given $\mathbf{h}^*$. But not all $|\mathbf{h}|$ are variationally integrable, so that the proof fails to go through in this way.

Ex. 34.1. If h is the function of Ex. 31.2 (p. 64) show that the ratio $h(I)/mI$ is bounded but that the derivative does not exist anywhere. Similarly for $mI/(mI+h(I))$. This suggests that for the conclusion of Theorem 34.3 to hold, both $\mathbf{h}$ and $\mathbf{h}^*$ must have a property near to finite additivity.

Ex. 34.2. Show that continuity of $\mathbf{h}^*$ in Theorem 34.3 can be dropped by using Theorem 32.2.

35. The differentiation of variational integrals

Instead of using previous theorems to obtain a partial result we prove the following directly.

NV

Th. 4.1.1

THEOREM 35.1. *If* K *is the indefinite variational integral of* $\{\mathrm{fh_l}, \mathrm{fh_r}\}$ *in an elementary set* E, *then*

$$\mathrm{D}(\mathrm{K}, \mathbf{h}; \mathrm{E}; \mathrm{x}) = \mathrm{f}(\mathrm{x}) \tag{35.1}$$

except possibly at points of a set X *with*

$$\mathrm{IV}(\mathbf{h}; \mathrm{E}; \mathrm{X}) = 0 \tag{35.2}$$

If $\mathrm{D}(\mathrm{K}, \mathbf{h}; \mathrm{E}; \mathrm{x})$ *exists at all points of a set* $\mathrm{X_1}$ *in* E *then* (35.1) *is true in* $\mathrm{X_1}$, *except possibly in a set* $\mathrm{X_2}$ *with*

$$\mathrm{V}(\mathbf{h}; \mathrm{E}; \mathrm{X_2}) = 0 \tag{35.3}$$

In cases when zero inner variation implies zero variation, (35.3) gives nothing fresh. But here we place no hypothesis on $\mathbf{h}$ save the existence of K.

If the derivative does not exist at x, or if it exists but is not equal to $f(x)$, then for some fixed $\epsilon_1 = \epsilon_1(x) > 0$ and each $\delta > 0$, there is either some $t < x$ such that $x-t < \delta$, $[t, x) \subseteq E$, and

$$|K(t, x)-f(x)h_l(t, x)| > \epsilon_1|h_l(t, x)| \tag{35.4}$$

or there is some $u > x$ such that $u-x < \delta$, $[x, u) \subseteq E$, and

$$|K(x, u)-f(x)h_r(x, u)| > \epsilon_1|h_r(x, u)| \tag{35.5}$$

Let X_{3n} be the set of such x with $\epsilon_1(x) \geqslant 1/n$. Further, given $\epsilon > 0$, there are a $\delta_1(x) > 0$ and a non-negative finitely super-additive interval function χ with $\chi(E) < \epsilon$, such that for $\delta \leqslant \delta_1(x)$, the left-hand side of (35.4) is bounded by $\chi(t, x)$, and the left-hand side of (35.5) is bounded by $\chi(x, u)$. Then either

$$|h_l(t. x)| \leqslant n\chi(t, x), \text{ or } |h_r(x, u)| \leqslant n\chi(x, u)$$

and the inner variation of X_{3n} is bounded by $n\epsilon$. As $\epsilon > 0$ is as small as we please, the inner variation is 0. This is true for $n = 1, 2, \ldots$, so that by Ex. 28.1 (p. 55) for IV we have (35.2). Note that the use of the inner variation is the simplest way of expressing the first differentiation result.

If $D(K, \mathbf{h}; E; x)$ exists in X_2, but is not $f(x)$, then for some $\epsilon_2 > 0$, $\delta_2 > 0$, both depending on x, every interval $[t, x)$ (respectively, $[x, u)$) with length less than δ_2 satisfies (35.4) (respectively (35.5)), with ϵ_2 for ϵ_1, so that (35.3) follows in a similar way. This result supplies another proof of (31.15) in Theorem 31.3 (p. 62).

Th. 4.1.2

THEOREM 35.2. *Let* $\mathbf{h}$, $\{\mathrm{fh_l}, \mathrm{fh_r}\}$, *and* $\{|\mathrm{fh_l}|, |\mathrm{fh_r}|\}$ *be variationally integrable in an elementary set* E, *with* $\mathrm{h_s} \geqslant 0$. *Then if* t *is the associated point of the general interval, and* H *is the integral of* $\mathbf{h}$,

$$\lim_{\mathrm{u}\to 0} \frac{1}{\underset{u\quad v}{\mathrm{H}(\mathrm{x}, \mathrm{x}+\mathrm{u})}} (\mathrm{V})\int_{[u,v)}^{\mathrm{u}} |\mathrm{f}(\overset{x}{\mathrm{x}}+\mathrm{t})-\mathrm{f}(\overset{u}{\mathrm{x}})|\, d\mathbf{h} = 0 \quad ([\mathrm{x}, \mathrm{x}+\mathrm{u}) \subseteq \mathrm{E})$$

except for x *in a set* $\mathrm{X_4}$ *satisfying* (35.2), *and similarly for* $\mathrm{u} < 0$.

The integrability of the three pairs of interval functions, and Theorems 19.1 (p. 27) and 25.1 (p. 43), give in turn, for each constant α, the variational integrability in E, with respect to $\mathbf{h}$, of

$$f(w)-\alpha, \text{ and } |f(w)-\alpha|$$

Thus by Theorem 35.1 we obtain

$$\lim_{u\to 0}\frac{1}{H(x, x+u)}(V)\int_0^u |f(x+t)-\alpha|\,d\mathbf{h}$$
$$= |f(x)-\alpha|([x, x+u) \subseteq E)$$

except for x in a set $X_5(\alpha)$ satisfying (35.2). Taking

$$\alpha = \pm m.2^{-n} \pm ip.2^{-q}\ (m, n, p, q = 0, 1, 2, \ldots)$$

with four combinations of signs, and using Theorem 28.1, the union X_4 of the corresponding $X_5(\alpha)$ also satisfies (35.2). Now let α be any constant, and let β be one of the special numbers within ϵ of α. Then

$$\left|(V)\int_0^u |f(x+t)-\alpha|\,d\mathbf{h}-(V)\int_0^u |f(x+t)-\beta|\,d\mathbf{h}\right|$$
$$\leqslant (V)\int_0^u \big||f(x+t)-\alpha|-|f(x+t)-\beta|\big|\,d\mathbf{h}$$
$$\leqslant |\alpha-\beta| H(x, x+u) \leqslant \epsilon H(x, x+u)$$

Also, for x not in X_4, and for u small enough,

$$(V)\int_0^u |f(x+t)-\beta|\,d\mathbf{h} \leqslant H(x, x+u)\{|f(x)-\beta|+\epsilon\}$$
$$\leqslant H(x, x+u)\{|f(x)-\alpha|+2\epsilon\}$$
$$(V)\int_0^u |f(x+t)-\alpha|\,d\mathbf{h} \leqslant H(x, x+u)\{|f(x)-\alpha|+3\epsilon\}$$

Here the α is arbitrary, and so can depend on x, so that if in particular we take $\alpha = f(x)$ we obtain the theorem.

Theorem† 35.3. *Let* $|\mathbf{h}|$ *and* $|\mathbf{h}|\mathrm{ch}(X_6)$ *be variationally integrable in an elementary set* E, *where* X_6 *is a set of* x. *If* $a \leqslant x \leqslant b$, $a < b$, *with* a, b *tending to* x, *then*

$$V(\mathbf{h}; [a, b]; X_6)/V(\mathbf{h}; [a, b]) \to \mathrm{ch}(X_6; x) \tag{35.6}$$

except for x *in a set satisfying Theorem* 35.1 *(35.2).*

If K is the variational integral of $|\mathbf{h}|$ then by Theorem 31.2 (31.7; 31.9) (pp. 60–1), (35.6) becomes

$$(V)\int_a^b ch(X_6)\,d|\mathbf{h}|/K(a, b) \to ch(X_6; x) \tag{35.7}$$

In (35.7) we can replace $|\mathbf{h}|$ by K, using Theorem 31.3 (31.20) (p. 63), so that by Theorem 35.1, (35.7) is true for $a = x$, for $b = x$, and so by finite additivity, for $a < x < b$, except for x in a set satisfying (35.2).

---

† The corresponding result in Lebesgue theory is the *density theorem*.

Assuming ∃ f, partial result due to Beppo Levi, "Sopra l'integrazione delle serie (Rend. del R. Ist. Lomb., 1906)
For lim ∫ see Ch. de la V Poussin, "Sur l'intégrale de Lebesgue, TAMS 1915
Vitali, "Sull' integrazione per serie (Rend. Circ. Palermo 1907)

LIMITS UNDER THE INTEGRAL SIGN

36. Monotone sequences and functions

It is well known that the Lebesgue integral handles with ease many problems concerning the commutation of the operations of integration and the taking of a limit. This ease also holds in our present theory, the proofs being simpler than ever. We begin with interval functions that are monotone increasing as some parameter tends to a 'limit'. Similar results hold when the functions are monotone decreasing. For simplicity we suppose that the parameter y tends to infinity, where either y takes all positive integer values (when we have a sequence), or all values $y \geqslant 0$. Cases in which $y \to a-$ or $y \to a+$ can be dealt with by obvious transformations.

Not all monotone increasing sequences of interval functions give reasonable results, see Ex. 20.1 and 20.2 (p. 29).

Gen^a.

THEOREM 36.1. *Let* $h_s \geqslant 0 (s = l, r)$, *and let the point function* $f(x, y)$ *be non-negative and monotone increasing in* y *for each fixed* x. *If for fixed* y, $f(x, y)$ *is variationally integrable with respect to* $\mathbf{h}$ *in an elementary set* E, *with indefinite variational integral* $K(., y)$, *and if* $K(E, y)$ *has a finite supremum as* y *tends to infinity then*

$$f(x) = \lim_{y\to\infty} f(x, y)$$

exists (finitely) except possibly in a set X *with*

$$V(\mathbf{h}; E; X) = 0 \tag{36.1}$$

Further, $f(x)$ *is variationally integrable with respect to* $\mathbf{h}$ *in* E, *with indefinite variational integral the limit*

$$K = \lim_{y\to\infty} K(., y)$$

so that

$$(V)\int_E \lim_{y\to\infty} f(., y)\,d\mathbf{h} = \lim_{y\to\infty} (V)\int_E f(., y)\,d\mathbf{h} \tag{36.2}$$

If $h_s \geqslant 0$ $(s = l, r)$, *if* $f(x, y) \geqslant 0$ *and monotone increasing in*

y for each fixed x, *if*

$$f(x) = \lim_{y\to\infty} f(x, y)$$

finite, except in X *satisfying* (36.1), *and if* $f(x, y)$ *is but* $f(x)$ *is not, variationally integrable with respect to* $\mathbf{h}$ *in* E, *then*

$$(V)\int_E f(., y)\,d\mathbf{h} \to +\infty \tag{36.3}$$

The case when y takes all values in $y \geqslant 0$, reduces to the sequence case, by the simple observation that

$$f(x, n) \leqslant f(x, y) \leqslant f(x, n+1) \quad (n < y < n+1)$$

For fixed j, $N > 0$, let X_{1j} be the set of x where $f(x, j) > N$, and let X_2 be the set where either

$$\lim_{j\to\infty} f(x, j) > N, \text{ or } f(x, j) \to \infty \text{ as } j \to \infty$$

Then by Ex. 26.6 (p. 48) and Theorems 31.2 (31.7) (p. 60), 30.1 (30.5) (p. 57) and 28.6 (p. 54), (6) M is the supremum of $K(E, j)$

$$M \geqslant (V)\int_E f(., j)\,d\mathbf{h} = V(\{f(., j)h_l, f(., j)h_r\}; E)$$

$$\geqslant V(\{f(., j)h_l, f(., j)h_r\}; E; X_{1j}) \geqslant N \,.\, V(\mathbf{h}; E; X_{1j}),$$

$$V(\mathbf{h}; E; X_{1j}) \leqslant M/N, \quad V(\mathbf{h}; E; X_2) \leqslant M/N$$

since X_2 is the limit of X_{1j}. Also, for all N, $X_2 \supseteq X$, so that if we let $N \to \infty$ and use Ex. 26.6 (p. 48) we have (36.1).

(p. 59) → Now $K(., j)$ is non-negative and finitely additive, and so is VB^*, so that by Theorem 30.4, $\mathbf{h}$ is VBG^* except possibly for the set where $f(x, j) = 0$. Hence by Theorem 29.1, there are a set $\mathbf{A}$ complete in E, a point function $k(x) > 0$, and a non-negative finitely superadditive interval function χ, such that (p. 56)

$$|h_s| \leqslant k(x)\chi, \text{ or } f(x, j) = 0 \text{ (all } j) \tag{36.4}$$

for intervals of $\mathscr{L} \cup \mathscr{R}$, where x is the associated point.

By Theorem 31.3, p. 62, (31.17) implying (31.16), we can replace $f(x, j)$ by 0 in X, so that we can assume that

$$f = \lim_{j\to\infty} f(., j)$$

everywhere. Thus, given $\epsilon > 0$, there are an integer $m(x)$, $\mathbf{A}_{1j}$

complete in E, and non-negative finitely superadditive interval functions χ_{1j}, such that

$$f(x)-f(x,j) < \epsilon/k(x) \ (j \geqslant m(x)), \text{ or } f(x) = 0 = f(x,j) \tag{36.5}$$

$$\chi_{1j}(E) < \epsilon.2^{-j} \tag{36.6}$$

$$|K(I,j)-f(x,j)h_s(I)| \leqslant \chi_{1j}(I) \tag{36.7}$$

for intervals in $\mathscr{L}_{1j} \cup \mathscr{R}_{1j}$ with associated point x. Further, for each fixed $x, f(x,j) \geqslant 0$, and is monotone increasing in j, so that by Theorems 20.1 (p. 28) and 21.1 (p. 31), $K(E_1, j)$ is monotone increasing in j for fixed elementary sets E_1 in E, and

$$K(E_1,j) \leqslant K(E,j) \leqslant M$$

and hence $K(E_1, j)$ tends to the limit $K(E_1)$. Thus for an integer J,

$$K(E)-K(E,j) < \epsilon \ (j \geqslant J) \tag{36.8}$$

It also follows that $K-K(.,j)$ is a non-negative interval function that is finitely additive, and so finitely superadditive.

As in Ex. 16.9 (p. 24) we form a complete set $\mathbf{A}_2$ by taking the $\mathbf{A}_6(x)$ there to be $\mathbf{A}.\mathbf{A}_{1j}$ for $j = \max(m(x), J)$. For intervals of $\mathscr{L}_2 \cup \mathscr{R}_2$ and this j, by (36.4; 36.5; 36.7),

$$\begin{aligned} f(x)h_s &< \{f(x,j)+\epsilon/k(x)\}h_s \\ &\leqslant K(.,j)+\chi_{1j}+\epsilon\chi \leqslant K+\sum_{j=1}^{\infty}\chi_{1j}+\epsilon\chi \end{aligned}$$

Also by (36.7),

$$\begin{aligned} f(x)h_s &\geqslant f(x,j)h_s \geqslant K(.,j)-\chi_{1j} \\ &\geqslant K-\{K-K(.,J)\}-\sum_{j=1}^{\infty}\chi_{1j} \end{aligned}$$

By (36.6; 36.8),

$$\sum_{j=1}^{\infty}\chi_{1j}(E)+\epsilon\chi(E)+K(E)-K(E,J) < \{2+\chi(E)\}\epsilon$$

so that K is variationally equivalent to $\{fh_l, fh_r\}$ in E, proving (36.2).

In the second part of the theorem, if (36.3) is false then $K(E, y)$ is bounded as $y \to \infty$, and $f(x)$ is variationally integrable with respect to $\mathbf{h}$ in E. As this last is given to be false, (36.3) must be true.

37. Majorized sequences and functions

In the theory of Lebesgue integration Fatou's Lemma is useful in showing that certain function spaces are complete. The corresponding result here is (37.3). Also (37.8; 37.9; 37.10) correspond to Lebesgue's majorized convergence result. L in Annales de l'École Normale, 1910.

✓ **Theorem 37.1.** *Let* $\mathrm{h}_s \geqslant 0 (s = l, r)$, *and let* $\mathrm{f}_1, \mathrm{f}_2, \mathrm{f}(., \mathrm{y})$ *be variationally integrable with respect to* $\mathbf{h}$ *in an elementary set* E, *with*

$$\mathrm{f}(\mathrm{x},\mathrm{y}) \geqslant \mathrm{f}_1(\mathrm{x}) \tag{37.1}$$

for each fixed y, *and, in the case when* y *takes all values in* $\mathrm{y} \geqslant 0$, *for each* $0 \leqslant \mathrm{Y} < \mathrm{Z}$ *let*

$$\inf_{\mathrm{Y} \leqslant \mathrm{y} \leqslant \mathrm{Z}} \mathrm{f}(\mathrm{x},\mathrm{y}) \tag{37.2}$$

be variationally integrable with respect to $\mathbf{h}$ *in* E. *Then*

$$(\mathrm{V})\int_{\mathrm{E}} \liminf_{\mathrm{y}\to\infty} \mathrm{f}(.,\mathrm{y})d\mathbf{h} \leqslant \liminf_{\mathrm{y}\to\infty} (\mathrm{V})\int_{\mathrm{E}} \mathrm{f}(.,\mathrm{y})d\mathbf{h} \tag{37.3}$$

If instead of (37.1) *and* (37.2) *we have, for each fixed* y,

$$\mathrm{f}(x, y) \leqslant \mathrm{f}_2(\mathrm{x}) \tag{37.4}$$

and, in the case when y *takes all values in* $\mathrm{y} \geqslant 0$, *for each* $0 \leqslant \mathrm{Y} < \mathrm{Z}$, *if*

$$\sup_{\mathrm{Y} \leqslant \mathrm{y} \leqslant \mathrm{Z}} \mathrm{f}(\mathrm{x},\mathrm{y}) \tag{37.5}$$

is variationally integrable with respect to $\mathbf{h}$ *in* E, *then*

$$(\mathrm{V})\int_{\mathrm{E}} \limsup_{\mathrm{y}\to\infty} \mathrm{f}(.,\mathrm{y})d\mathbf{h} \geqslant \limsup_{\mathrm{y}\to\infty} (\mathrm{V})\int_{\mathrm{E}} \mathrm{f}(.,\mathrm{y})d\mathbf{h} \tag{37.6}$$

If in (37.3) *or* (37.6),

$$\lim_{\mathrm{y}\to\infty} \mathrm{f}(\mathrm{x},\mathrm{y}) \tag{37.7}$$

exists, then we do not need the variational integrability of (37.2) *or* (37.5) *with respect to* **h** *in* E.

If

$$\mathrm{f}_1(\mathrm{x}) \leqslant \mathrm{f}(\mathrm{x}, \mathrm{y}) \leqslant \mathrm{f}_2(\mathrm{x}), \ \mathrm{f}(\mathrm{x}) = \lim_{\mathrm{y}\to\infty} \mathrm{f}(\mathrm{x}, \mathrm{y}) \tag{37.8}$$

except possibly in a set X *of* x *with*

$$\mathrm{V}(\mathbf{h}; \mathrm{E}, \mathrm{X}) = 0 \tag{37.9}$$

then if we put f(x) = 0 *for* x *in* X,

$$(\mathrm{V}) \int_{\mathrm{E}} \mathrm{f}\, d\mathbf{h} = \lim_{\mathrm{y}\to\infty} (\mathrm{V}) \int_{\mathrm{E}} \mathrm{f}(., \mathrm{y})\, d\mathbf{h} \tag{37.10}$$

In particular, if **h** *is also variationally integrable in* E, *the results are true when* $\mathrm{f}_1, \mathrm{f}_2$ *are constants, so that* f(x, y) *is bounded above, or below, or both, in* x *and* y. (37.1

As by Theorem 19.1 (p. 27) the variational integral is a distributive functional we can subtract f_1 from $f(., y)$ and in (37.1) assume that $f(x, y) \geqslant 0$. First taking the case when y has positive integer values, we have

$$f(., j) \leqslant \sum_{k=m}^{n} f(., k) \ (m \leqslant j \leqslant n)$$

so that by Theorem 25.1 (p. 43) applied to $-f(., j)h_s$, for $m \leqslant j \leqslant n$,

$$\min_{m \leqslant j \leqslant n} f(., j) \tag{37.12}$$

is variationally integrable with respect to **h** in E. This corresponds to (37.2), so that we can now consider both cases together. The function (37.2) is monotone decreasing as Z tends to infinity, and is bounded below by 0, so that

$$f(x, Y) - \inf_{Y \leqslant y \leqslant Z} f(x, y) \geqslant 0$$

is monotone increasing in Z, and is bounded above by $f(x, Y)$. Applying Theorem 36.1 p. 82 we see that by distributivity,

$$\inf_{y \geqslant Y} f(x, y) \tag{37.13}$$

is variationally integrable with respect to **h** in E, with

$$(V) \int_E \inf_{y \geqslant Y} f(., y)\, \mathrm{d}\mathbf{h}$$

$$= \lim_{Z\to\infty} (V) \int_E \inf_{Y \leqslant y \leqslant Z} f(., y)\, \mathrm{d}\mathbf{h}$$

$$\leqslant \inf_{y \geqslant Y} (V) \int_E f(., y)\, \mathrm{d}\mathbf{h} \tag{37.14}$$

Again (37.13) is monotone increasing as Y tends to infinity; and we can obviously assume that the right-hand side of (37.3) is finite, so that by (37.14) the variational integral of (37.13) with respect to **h** is bounded as Y tends to infinity. Hence by Theorem 36.1, first part, p. 82,

$$\liminf_{y\to\infty} f(x, y) = \lim_{Y\to\infty} \{ \inf_{y \geqslant Y} f(x, y)\} \tag{37.15}$$

is finite, except possibly in a set X_1 satisfying (37.9), and Theorem 36.1 and (37.14; 37.15) give (37.3). To prove (37.6) from (37.4; 37.5) we replace $f(x, y)$ by

$$f_2(x) - f(x, y)$$

If, however, (37.7) exists, except possibly in a set X_2 satisfying (37.9), then the limit is

$$\lim_{n\to\infty} f(x, y_n)$$

so that without (37.2) or (37.5), we have in (37.3) for a suitable sequence $\{y_n\}$ tending to infinity,

$$\liminf_{y\to\infty} (V) \int_E f(., y)\, \mathrm{d}\mathbf{h}$$

$$= \lim_{n\to\infty} (V) \int_E f(., y_n)\, \mathrm{d}\mathbf{h} \geqslant (V) \int_E \lim_{n\to\infty} f(., y_n)\, \mathrm{d}\mathbf{h}$$

$$= (V) \int_E \lim_{y\to\infty} f(., y)\, \mathrm{d}\mathbf{h}$$

and similarly for (37.6). Thus without (37.2) or (37.5), we obtain (37.10) from (37.3; 37.6; 37.8; 37.9).

Note that Ex. 37.1 shows that the variational integrability of (37.2; 37.5) with respect to **h** in E, is vital to the theorem, except in the cases when the limit exists or y takes only a countable number of values.

We can obtain information about the continuity in y of the integral of $f(x, y)$ by restating part of Theorem 37.1 using $y \to a$.

✓ **Theorem 37.2.** *If*

$$\lim_{y\to a} f(x, y) = f(x, a)$$

except possibly for x *in a set satisfying* (37.9), *if* $f_1, f_2, f(., y)$ *are all variationally integrable with respect to* **h** *in the elementary set* E, *with* $h_s \geqslant 0$ (s = 1, r) *and*

$$f_1 \leqslant f(., y) \leqslant f_2 \ (a-\delta \leqslant y \leqslant a+\delta, \textit{for some } \delta > 0)$$

then

$$\lim_{y\to a} (V) \int_E f(., y)\, d\mathbf{h} = (V) \int_E f(., a)\, d\mathbf{h}$$

Similarly we can deal with differentiation under the integral sign.

✓ **Theorem 37.3.** *If* $\partial f(x, y)/\partial y$ *exists at* y = a *for each fixed* x *in an elementary set* E, *except possibly in a set satisfying* (37.9), *and if* $f_1, f_2, f(., y)$ *are all variationally integrable with respect to* **h** *in* E, *with* $h_s \geqslant 0$ (s = 1, r), *and*

$$f_1 \leqslant \{f(., y)-f(., a)\}/(y-a) \leqslant f_2$$

$$(0 < |y-a| \leqslant \delta, \textit{for some } \delta > 0), \textit{ then}$$

$$\frac{d}{dy}(V) \int_E f(., y)\, d\mathbf{h} = (V) \int_E \frac{\partial f(., y)}{\partial y} d\mathbf{h}$$

A more general case where we differentiate with respect to interval functions on the y-axis, can be dealt with in a similar way.

Ex. 37.1. The conditions (37.2; 37.5) of Theorem 37.1 are necessary if y takes all values in $y \geqslant 0$ unless the limit of $f(., y)$ exists a.e. For let X_1 be a set for which

$$(V) \int_E ch(X_1)\, dx$$

does not exist. A discussion on such sets will be given in Ex. 38.1. For $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$ let us put

$$f(x, y) = \begin{cases} 1 (x = y \in X_1), \\ 0 (\textit{otherwise}); \end{cases} \quad f(x, y+n) = f(x, y) \ (n = 1, 2, \ldots)$$

Then f is bounded in $0 \leqslant x \leqslant 1$, $y \geqslant 0$, while

$$\limsup_{y\to\infty} f(x, y) = \sup_{0\leqslant y\leqslant 1} f(x, y)$$

is the characteristic function of X_1, and so is not variationally integrable with respect to x in [0, 1].

38. Absolute integration

The absolute case has many features in common with Lebesgue theory. For example, the following gives useful results.

✓ **Theorem 38.1.** *Let* $h_s \geqslant 0$ (s = 1, r), *and let* f_j, $|f_j|$ *be variationally integrable with respect to* **h** *in an elementary set* E, *for* j = 1, 2, ... *If*

$$\sum_{j=1}^{\infty} (V) \int_E |f_j|\, d\mathbf{h} \tag{38.1}$$

is finite then

$$f = \sum_{j=1}^{\infty} f_j$$

exists as an absolutely convergent series, except possibly in a set X *with*

$$V(\mathbf{h}; E; X) = 0$$

and f *is variationally integrable with respect to* **h** *in* E, *with*

$$(V) \int_E f\, d\mathbf{h} = \sum_{j=1}^{\infty} (V) \int_E f_j\, d\mathbf{h} \tag{38.2}$$

Further, the variational integral of $f\,ch(X_1)$ *with respect to* **h** *in* E *is completely additive in* X_1 *whenever the corresponding integrals using* $|f|ch(X_1)$ *exist. More precisely, let* $\{X_{2j}\}$ *be a sequence of mutually exclusive sets with union* X_1. *If, for* j = 1, 2, ..., $f\,ch(X_{2j})$ *and* $|f|ch(X_{2j})$ *are both variationally* *integrable with respect to* $\mathbf{h}$ *with* $h_s \geqslant 0$ $(s = 1, r)$, *and if*

$$\sum_{j=1}^{\infty} (V) \int_E |f| ch(X_{2j})\, d\mathbf{h} \tag{38.3}$$

is convergent, then there exists

$$(V) \int_E f \,.\, ch(X_1)\, d\mathbf{h} = \sum_{j=1}^{\infty} (V) \int_E f \,.\, ch(X_{2j})\, d\mathbf{h} \tag{38.4}$$

and similarly for $|f|$ *replacing* f.

Result (38.4) follows from (38.3) because (38.2) follows from (38.1). To prove the latter implication we use the distributivity of the integral,

$$(V) \int_E \sum_{j=1}^{k} |f_j|\, d\mathbf{h} = \sum_{j=1}^{k} (V) \int_E |f_j|\, d\mathbf{h} \leqslant \sum_{j=1}^{\infty} (V) \int_E |f_j|\, d\mathbf{h}$$

which is convergent, so that by Theorem 36.1, p. 82,

$$f_0 = \sum_{j=1}^{\infty} |f_j|$$

is convergent except in X, and f exists, and f_0 is variationally integrable in E with respect to $\mathbf{h}$. In Theorem 37.1 we now replace $f_1, f_2, f(., k)$ by the present

$$-f_0,\ f_0,\ \sum_{j=1}^{k} f_j$$

and obtain (38.2) from (37.10).

For a real point function f and constants $a < b$ let

$$X(f = a),\ X(f > a),\ X(f \geqslant a),\ X(f < a),$$
$$X(f \leqslant a),\ X(a \leqslant f < b), \ldots$$

denote the sets of points x for which the appropriate equality or inequalities hold.

THEOREM 38.2. *Let* $\mathbf{h}$ *be variationally integrable in an elementary set* E *with* $h_s \geqslant 0$ $(s = 1, r)$. *Let* f *and* $|f|$ *be real and variationally integrable in* E *with respect to* $\mathbf{h}$. *For a real constant* b *let* $f_0{}^b$ *be the characteristic function of* $X(f \geqslant b)$.

Then $f_0{}^b$ *is variationally integrable with respect to* $\mathbf{h}$ *in* E *with variational integral*

$$V(\mathbf{h};\, E;\, X(f \geqslant b)) \tag{38.5}$$

Similar results hold for the sets $X(f = a), X(f > a), \ldots$

From Theorem 25.4 (p. 46), $\max(f, a)$ and $\min(f, a)$ are variationally integrable with respect to $\mathbf{h}$ in E, and hence for constants $a < b$, so is

$$f_0{}^{ab} = \max(\min(f, b), a)$$

This is f when $a \leqslant f \leqslant b$, it is a when $f \leqslant a$, and is b when $f \geqslant b$. For fixed b,

$$f_0{}^b = \lim_{a \to b-} (f_0{}^{ab} - a)/(b - a)$$

stet ✓

the function inside the limit ~~being monotone decreasing and~~ lying between 0 and 1. Hence by Theorem 37.1 (37.10), p. 86, $f_0{}^b$ is variationally integrable with respect to $\mathbf{h}$ in E. Then by Theorem 31.2 (31.9), p. 61, we have (38.5).

The characteristic function of $X(f > a)$ is the limit of $f_0{}^b$ as $b \to a+$, so that by Theorems 37.1 (37.10) p. 85 and 31.2 (31.9) p. 61 we have a similar result. Also

$$ch(X(f = a)) = \min\big(ch(X(f \geqslant a), 1 - ch(X(f > a))\big)$$
$$ch(X(f \leqslant a)) = 1 - ch(X(f > a)),$$
$$ch(X(f < a)) = 1 - ch(X(f \geqslant a)),$$
$$ch(X(a \leqslant f < b)) = \min\big(ch(X(f \geqslant a)), ch(X(f < b))\big), \ldots$$

so that the rest of the theorem follows.

Ex. 38.1. Given x in $[0, 1)$, let $X(x)$ be the set of all y in $[0, 1)$ such that $y - x$ is rational. Let M be a set in $[0, 1)$ such that for each x in $[0, 1)$, $M \cap X(x)$ consists of exactly one point. Show that $ch(M)$ is not variationally integrable with respect to x in $[0, 1]$.

(Let the rationals in $[0, 1)$ be put as a sequence $\{r_n\}$, and let M_n be the set of all $x + r_n$, $x + r_n - 1$, that lie in $[0, 1)$ and have x in M. Then

$$(V) \int_0^1 ch(M_n)\, dx = (V) \int_0^1 ch(M)\, dx \tag{38.6}$$

if the second integral exists. Hence by distributivity, for all n,

$$n(V)\int_0^1 ch(M)\,\mathrm{d}x = \sum_{j=1}^{n}(V)\int_0^1 ch(M_j)\,\mathrm{d}x$$

$$= (V)\int_0^1 ch\Big(\bigcup_{j=1}^{n} M_j\Big)\,\mathrm{d}x \leqslant 1,\ (V)\int_0^1 ch(M)\,\mathrm{d}x = 0$$

This with (38.6), Theorem 38.1 (38.3), and

$$\bigcup_{j=1}^{\infty} M_j = [0, 1)$$

gives a contradiction, so that the integrals in (38.6) cannot exist.)

39. Limits of step functions

By using sequences of functions constant in a finite number of intervals we obtain two interesting results.

THEOREM 39.1. *Let* $\mathbf{h}, |\mathbf{h}|, \mathbf{h}^*$ *be variationally integrable,* $\mathbf{h}^*$ *continuous, and* $\mathrm{h}_s^* \geqslant 0(\mathrm{s} = 1, \mathrm{r})$, *all in an elementary set* E. *Then* $\mathrm{f(x)} = \mathrm{D}(\mathbf{h}, \mathbf{h}^*; \mathrm{E}; \mathrm{x})$ *exists in* E, *except for a set* X *of* x *satisfying*

$$\mathrm{V}(\mathbf{h}^*; \mathrm{E}; \mathrm{X}) = 0 \tag{39.1}$$

and if we put f = 0 *in* X, *then* f *and* |f| *are variationally integrable with respect to* $\mathbf{h}^*$ *in* E, *with*

$$(\mathrm{V})\int_{\mathrm{E}} |\mathrm{f}|\,d\mathbf{h}^* \leqslant \mathrm{V}(\mathbf{h}; \mathrm{E}) \tag{39.2}$$

The existence of the derivative, except in X satisfying (39.1), is proved in Theorem 34.3 (p. 76). We have now to prove the rest. By Theorem 34.2 (p. 75) we can replace $\mathbf{h}, \mathbf{h}^*$ by their variational integrals, respectively, H and H^*. Then we can split up H into its positive and negative parts, as in Theorem 31.1 (31.5), p. 60, so that we need only assume that

$$h_l = h_r = H \geqslant 0,\ h_l^* = h_r^* = H^* \geqslant 0$$

where H, H^* are finitely additive. By Theorems 21.1(p. 31) and 28.2 (p. 50) we can also assume that $E = [a, b]$, the more general case being obtained by finite union. Let us put

$$a_{jk} = a+(b-a)j.2^{-k}(j = 0, 1, 2, \ldots, 2^k; k = 1, 2, \ldots)$$

$$f_k(x) = H(a_{j-1,k}, a_{jk})/H^*(a_{j-1,k}, a_{jk})(a_{j-1,k} \leqslant x < a_{jk},$$
$$j = 1, 2, \ldots, 2^k), f_k(b) = 0$$

If the denominator of the fraction is 0 we put $f_k(x) = 0$; the interval then lies in X. Then by continuity of H^* at the points a_{jk} we obtain

$$(V)\int_a^b f_k\,\mathrm{d}H^* = \sum_{j=1}^{2^k} f_k(a_{j-1,k})H^*(a_{j-1,k}, a_{jk})$$

$$= \sum_{j=1}^{2^k} H(a_{j-1,k}, a_{jk}) = H(a, b) \tag{39.3}$$

Further, for fixed x not in X, with $a_{j-1,k} < x < a_{jk}$, and given $\epsilon > 0$, then by the definition (34.1), p. 74, of differentiation there is an integer k_0 such that

$$x-\delta < a_{j-1,k} < a_{jk} < x+\delta(k \geqslant k_0),$$
$$|H(a_{j-1,k}, x)-f(x)H^*(a_{j-1,k})| \leqslant \epsilon H^*(a_{j-1,k}, x)$$
$$|H(x, a_{jk})-f(x)H^*(x, a_{jk})| \leqslant \epsilon H^*(x, a_{jk})$$

so that by finite additivity,

$$|H(a_{j-1,k}, a_{jk})-f(x)H^*(a_{j-1,k}, a_{jk})|$$
$$\leqslant \epsilon H^*(a_{j-1,k}, a_{jk}), f_k(x) \to f(x)$$

As H^* is continuous its variation at a single point, relative to E, is zero, so that by Theorem 28.1 (p. 49) its variation is zero in the countable set of points a_{jk}, relative to E, and with (39.1) we see that $f_k(x) \to f(x)$ except for x in a set satisfying (39.1). As $f_k \geqslant 0$, (39.3) and Theorem 37.1 (37.3) p. 85 give

$$(V)\int_a^b f\mathrm{d}H^* = (V)\int_a^b \lim_{k\to\infty} f_k\,\mathrm{d}H^*$$

$$\leqslant \liminf_{k\to\infty}(V)\int_a^b f_k\,\mathrm{d}H^* = H(a, b)$$

By Theorem 31.1 (31.3), p. 60, this is (39.2) in the particular case, so that we have completed the proof.

Theorem 39.2. *Let* $\mathbf{h}$ *be variationally integrable in an elementary set* E *with variational integral* H, *such that*

$$h_s \geqslant 0 \ (s = l, r).$$

Let $\{fh_l, fh_r\}$ *be variationally integrable in* E. *Then there is a sequence* $\{f_k\}$ *of step functions such that*

$$\lim_{k\to\infty} f_k = f \tag{39.4}$$

except possibly in a set X_1 *with*

$$IV(\mathbf{h}; E; X_1) = 0 \tag{39.5}$$

Can replace IV by V
See Lee JLMS 40 (1965) [illegible]8-344 (p. 340 Lemma 3)

and

$$(V)\int_E f_k\,d\mathbf{h} = (V)\int_E f\,d\mathbf{h} \tag{39.6}$$

If H(x−) *and* H(x+) *are only zero together, then in* (39.5) IV *becomes* V.

By Theorems 30.1 (30.2), p. 57, and 31.3 (31.20), p. 63, we can assume that $h_l = h_r = H$, finitely additive, and non-negative. Then if $H(u, v) = 0$, we have $H(u', v') = 0$ for all $(u', v') \subseteq (u, v)$, so that

$$(V)\int_u^v f\,dH = 0$$

also. Further, H is continuous except at a countable set at most. Thus we can use a construction similar to that used in Theorem 39.1, again taking the elementary set to be an interval $[a, b]$. This time the a_{jk}, for j in $0 \leqslant j \leqslant j_k$, and $k = 1, 2, \ldots$, are to be points of continuity of H, such that

$$\max(a_{jk} - a_{j-1,k}) \to 0, \quad a = a_{0k} < a_{1k} < \ldots < a_{\boxed{k}k} = b$$

j_k

as $k \to \infty$. Then H, H^* are replaced by

$$(V)\int f\,dH, H$$

and a similar proof applies. Naturally, instead of Theorem (p. 78) → 34.3 we use Theorem 35.1. We obtain (39.4; 39.5; 39.6), and then have no need to proceed to 'Fatou's lemma' (Theorem 37.1 (37.3)). We have to use the inner variation, unless the discontinuities $H(x-)$, $H(x+)$ are only zero together, when we can use Theorem 32.3 (p. 68). False!

A function f that is the limit of a sequence $\{f_n\}$ of step functions, except possibly in a set in which the variation of $\mathbf{h}$ relative to E, is zero, can be called $\mathbf{h}$-*measurable*. What we have proved is that *if* $h_s \geqslant 0(s = l, r)$, *with* $\mathbf{h}$ *variationally integrable in* E, *and with*

$$\lim_{t\to x-} h_l(t, x) \text{ and } \lim_{u\to x+} h_r(x, u)$$

both zero or both non-zero, and if f *is variationally integrable with respect to* $\mathbf{h}$ *in* E, *then* f *is* $\mathbf{h}$-*measurable.* If the sequence $\{f_n\}$ is bounded, or limited by two functions each variationally integrable with respect to $\mathbf{h}$ in E, then the converse is true. But not all unbounded sequences $\{f_n\}$ are limited, see Ex. 39.1.

If f is real and $\mathbf{h}$-measurable, then, except for the set of variation zero,

$$ch(X(f > a)) = \lim_{n\to\infty} ch(X(f_n > a))$$

for each real constant a, so that for the given $\mathbf{h}$,

$$(V)\int_E ch(X(f > a))\,d\mathbf{h} = \lim_{n\to\infty} (V)\int_E ch(X(f_n > a))\,d\mathbf{h}$$

and the variational integrability of $ch(X(f > a))$ with respect to $\mathbf{h}$ follows from that of $ch(X(f_n > a))$, which is constant in intervals.

Conversely, if $ch(X(f > a))$ is variationally integrable with respect to $\mathbf{h}$ in E, for each real constant a, we put

$$b_{jk} = j\,.\,2^{-k} \ (j = 0, \pm 1, \pm 2, \ldots, k = 0, 1, 2, \ldots),$$

$$f_k(x) = \sum_{j=-\infty}^{\infty} b_{jk}\,ch(X(b_{jk} \leqslant f < b_{j+1,k}); x), \ \lim_{k\to\infty} f_k = f$$

Thus f is the double limit of finite sums of multiples of characteristic functions that are variationally integrable with respect to $\mathbf{h}$ in E, so that f is the triple limit of step functions, apart from sets of variation zero. To show that f is a single limit of step functions apart from a set of variation zero, and so to prove the converse, we use the following theorem. It follows that we could have defined a real function f to be $\mathbf{h}$-measurable if for each real constant a, $ch(X(f > a))$ is variationally integrable with respect to $\mathbf{h}$ in E.

Leb's orig def is $\int = \lim \sum_{j=1}^{n} b_j \text{ meas } E\left(b_j \leqslant f < b_{j+1}\right)$
$b_1 = 0$, $b_n = \sup f$.

Theorem 39.3. *If the real* $\mathrm{f}_j (j = 1, 2, \ldots)$ *are* **h**-*measurable, and if* α *is a real constant, then*

$$\alpha \mathrm{f}_1,\ \mathrm{f}_1+\mathrm{f}_2,\ \mathrm{f}_1\mathrm{f}_2,\ \min(\mathrm{f}_1, \mathrm{f}_2),\ \max(\mathrm{f}_1, \mathrm{f}_2),\ \inf_{j \geqslant 1} \mathrm{f}_j,$$
$$\sup_{j \geqslant 1} \mathrm{f}_j,\ \liminf_{j\to\infty} \mathrm{f}_j,\ \limsup_{j\to\infty} \mathrm{f}_j,\ \lim_{j\to\infty} \mathrm{f}_j$$

are **h**-*measurable, where the limit need only exist except for a set of variation zero.*

By Theorem 28.1 (p. 49), the union of two sets of variation zero is also of variation zero. This is all that we need to prove the first five results. To prove the last five we prove the last result first, as follows.

$$\left.\begin{aligned} &\textit{If } \mathrm{f} = \lim_{j\to\infty} \mathrm{f}(., j) \textit{ except in a set of variation zero,}\\ &\textit{and if each } \mathrm{f}(., j) \textit{ is } \mathbf{h}\textit{-measurable, so is } \mathrm{f}. \end{aligned}\right\} \quad (39.7)$$

By Theorem 28.1 the exceptional sets for the separate $f(., j)$ have a union that is a set of variation zero. Outside this set we have

$$f(., j) = \lim_{k\to\infty} f(., j, k),\ \lim_{k\to\infty} \sup_{m>k} |f(., j, m) - f(., j, k)| = 0$$

p.54, so that by Theorem 28.6 there is an integer k_n such that, for all $k \geqslant k_n$,

$$V(\mathbf{h}; E; X(\sup_{m>k} |f(., j, m) - f(., j, k)| \leqslant 2^{-n})) > V(\mathbf{h}; E) - 2^{-n}$$

We can now renumber the functions $f(., j, k)$ so that we can replace k_n by n, and then, as $f(., j, n)$ is a step function, and so constant in intervals, we have for $j = 1, 2, \ldots$

$$V(\mathbf{h}; E; X(|f(., j, n+1) - f(., j, n)| \leqslant 2^{-n})) > V(\mathbf{h}; E) - 2^{-n}$$
$$V\big(\mathbf{h}; E; X(|f(., j, n+1) - f(., j, n)| > 2^{-n})\big) < 2^{-n}$$
$$V\big(\mathbf{h}; E; X(|f(., j) - f(., j, n)| > 2^{1-n})\big)$$
$$\leqslant V\Big(\mathbf{h}; E; \bigcup_{\nu=n}^{\infty} X(|f(., j, k+1) - f(., j, k)| > 2^{-k})\Big) < 2^{1-n}$$

[illegible]

Hence, in particular,

$$V\big(\mathbf{h}; E; X(|f(., n) - f(., n, n)| > 2^{1-n})\big) < 2^{1-n}$$
$$V\big(\mathbf{h}; E; X(|f(., n) - f(., n, n)| > 2^{1-n} \text{ for some } n \geqslant k)\big) < 2^{2-k}$$

In the rest of E we have

$$|f(., n) - f(., n. n)| \leqslant 2^{-n} \text{ (all } n \geqslant k)$$

so that where

$$f = \lim_{n\to\infty} f(., n)$$

we have

$$f = \lim_{n\to\infty} f(., n, n)$$

This is therefore true except in a set of variation less than 2^{2-k}. As k is an arbitrary integer the set has variation zero, and we have proved (39.7). It follows that

$$\inf_{j\geqslant 1} f_j = \lim_{k\to\infty} \min_{1\leqslant j\leqslant k} f_j,\ \sup_{j\geqslant 1} f_j = -\inf_{j\geqslant 1}(-f_j),\ \liminf_{j\to\infty} f_j$$
$$= \lim_{k\to\infty} \inf_{j\geqslant k} f_j,\ \limsup_{j\to\infty} f_j = \lim_{k\to\infty} \sup_{j\geqslant k} f_j$$

are also **h**-measurable, under the hypotheses of the theorem.

Thus the usual processes of analysis, when applied to **h**-measurable functions, give **h**-measurable functions.

Ex. 39.1. For $0 < x < 1$ let $f(x) = y.\log^{-3/2} y\, (y = x^{-1})$, while $f(0) = 0$. Show that

$$(V) \int_t^{\frac{1}{2}} f\, dx = 2(\log 2)^{-\frac{1}{2}} - 2(\log t^{-1})^{-\frac{1}{2}} \to 2(\log 2)^{-\frac{1}{2}}$$

as $t \to 0+$. By Theorem 46.1 (p. 113), the variational integral exists for $[0, \frac{1}{2}]$ with value the limit. On the other hand, in Theorem 39.2 we take $a_{1k} = 1/k$, so that

$$f_k(x) = k(V) \int_0^{1/k} f dx = 2k(\log k)^{-1/2} \quad (0 \leqslant x < 1/k)$$

If g majorizes the sequence $\{f_k\}$ show that g can have no variational integral with respect to x in $[0, \frac{1}{2}]$, using

$$(V) \int_t^{\frac{1}{2}} \frac{2}{x}\left(\log \frac{1}{x}\right)^{-\frac{1}{2}} dx \to \infty$$

as $t \to 0+$. Thus the sequence $\{f_k\}$ of step functions of Theorem 39.2 is not always majorized by a function that is variationally integrable with respect to **h** in E.

Ex. 39.2. Show that a continuous real function f is **h**-measurable, so that f is variationally integrable in E with respect to **h**, if **h** is VB^* and variationally integrable in E.

Ex. 39.3. If f is VB^* and g is continuous, show that f is variationally integrable with respect to g. (Use Ex. 39.2 and 33.2 (p. 73) and integrate by parts.)

Ex. 39.4. By using Ex. 34.2 p.78 show that continuity of **h*** can be dropped in Theorem 39.1. (Use a construction as in Theorem 39.2.)

40. An absolute integral is *AC**, and the integrability of a product

We can now improve Theorem 30.3 p59 in the case when the integral is absolute, giving a converse to Theorem 34.5 p.77

Theorem 40.1.† *If* h_s *is real for* $s = 1, r$, *and if* $|\mathbf{h}|$, $\{fh_1, fh_r\}$, *and* $\{|fh_1|, |fh_r|\}$ *are all variationally integrable in the elementary set* E, *then* Th. 3.2.5

here we need not have h ≥ 0.

$$K = (V)\int f\,d\mathbf{h}$$

is AC* *with respect to* **h** *in* E, $\boxed{\textit{and VB* in E}}$

By Theorem 19.2 (p. 27) we can suppose that f is real, and then by Theorem 25.4, (p.46)

$$f_M = \max(M, |f|) - M$$

is variationally integrable with respect to $|\mathbf{h}|$ in E. Let X_M be the set where $f_M \neq 0$. Then

$$f_M = (|f| - M)ch(X_M) \to 0$$

as $M \to \infty$, and is bounded by $|f|$, so that by Theorem 37.1 (37.8; 37.10), p. 86,

$$K_1 = (V)\int f_M\,d|\mathbf{h}| \to 0$$

as $M \to \infty$. Given $\epsilon > 0$, let M be so large that $K_1(E) < \tfrac{1}{2}\epsilon$, and then X so small that $V(\mathbf{h}; E; X) < \epsilon/(2M)$. Then if

$$f_M^* = |f| - f_M,\ |f_M^*| \leqslant M,\ K_2 = (V)\int f_M^*\,d|\mathbf{h}|$$

---

† This corresponds to Lebesgue theory as contrasted with Denjoy theory.

we have by Theorems 31.2 (31.12) (p.61), 31.1 (31.3) (p.60), 30.1 (30.3), (p.57) and Ex. 26.6 (p.48)

$$V(K; E; X) \leqslant V(K_1; E) + V(K_2; E; X) \leqslant \tfrac{1}{2}\epsilon + M.V(\mathbf{h}; E; X) < \epsilon$$

This proves the first part of the theorem. The second part is obtained by taking $X = E$ and omitting the final $< \epsilon$, as soon as it is noted that **h** is VB^* from Theorem 31.2 (31.7) (p.60).

Theorem 40.2. *Let* f *be variationally integrable with respect to a continuous* **h** *in an elementary set* E, *and let* g *be of bounded variation. Then*

Vital in Sargent check.

$$(V)\int_E fg\,d\mathbf{h} \tag{40.1}$$

exists. If $\{fh_1, fh_r\}$, $\{|fh_1|, |fh_r|\}$, $|\mathbf{h}|$ *are all integrable in* E, *where* $h_s (s = 1, r)$ *are real, and if* g *is bounded and* **h**-*measurable, then again* (40.1) *exists.*

In the first case,

$$(V)\int f\,d\mathbf{h} \tag{40.2}$$

is continuous, by Theorem 21.2 (21.12; 21.13) (p.33), so that by Ex. 39.3 (p.78), g is variationally integrable with respect to (40.2) in E. In the second case (40.2) is finitely additive and VB^* and so is the difference of two finitely additive non-negative interval functions. Thus g is variationally integrable with respect to (40.2) in E. We have used Theorems 21.1 (p.31), 40.1 [illegible]; 31.1 (31.5); (p.60) and 37.1 (p.85). In both cases the last step, giving (40.1), is obtained from Theorem 31.3 (31.20) (p.63).

These results are more precise than the third result of Theorem 39.3 (p.96).

Lee (21-12-65) mentions that analogue of Tonelli's Th. is true.
[See E.J. McShane, Integration, p. 145]
Fubini, "Sugli integrali multipli" (Rend. del R. Acc. d. Lincei, 1er semestre 1907)

6

DOUBLE INTEGRALS AND FUBINI'S THEOREM

41. The plane

We can define Riemann-complete and variational integrals in n-dimensional space, or even in an abstract space, as in Henstock (1961a and b). However, too great a generality might obscure the essential simplicity of the present approach to integration, so that here I will be content with extending the theory to integrals in a plane.

Instead of intervals we use rectangles R_w with sides parallel to the x- and y-axes of the rectangular Cartesian coordinate system, and all rectangles mentioned in this book are to be of this type. The x-axis is taken to be horizontal and the y-axis vertical, and we put $w = (x, y)$ and $R_w = (I_x, I_y)$, where I_x, I_y are the projections of R_w on the x- and y-axes, respectively. Each R_w has four vertices, that can be numbered clockwise from the top left-hand vertex and used as *associated points*.

For divisions of closed intervals we have used intervals open on the right. Here, a division of a closed rectangle consists of half-open rectangles. More precisely, given four real numbers $a < b, u < v$, we define the *closed rectangle* $R_w = ([a, b], [u, v])$ to be the set of points (x, y) for all x, y satisfying $a \leqslant x \leqslant b, u \leqslant y \leqslant v$. *The half-open rectangle* $R_w^- = ([a, b),[u, v))$ is the set of points (x, y) satisfying $a \leqslant x < b, u \leqslant y < v$. Thus if R_w is closed, I_x and I_y are closed intervals, while if R_w^- is half-open, I_x is open on the right, and so is I_y if we rotate the y-axis bodily through a quarter turn to bring the part with $y \geqslant 0$ into coincidence with the part $x \geqslant 0$ of the x-axis. Thus we can define the Riemann-complete and variational integrals using the I_x, I_y. *An elementary set* B_w is the union of a finite number of closed rectangles. Clearly, these can be taken as non-overlapping, where two rectangles are *overlapping* if they have a rectangle in common.

A division $\mathscr{D}$ of the elementary set B_w is a finite family of disjoint half-open rectangles R_{w1}^-, such that the union of the corresponding closed rectangles is B_w. To exhibit the R_{w1}^- we sometimes write $\mathscr{D}\{R_{w1}^-\}$. We also write B_w^- for the union of the R_{w1}^-. With the division we give a specification stating which of the vertices is to be the associated point of each rectangle in $\mathscr{D}$.

For $j = 1, 2, 3, 4$, a family $\mathscr{E}^j$ of half-open rectangles is *j-complete in an elementary set B_w*, if to each point w_1, that is the jth vertex of some rectangle in B_w, there corresponds a half-open rectangle R_{w2} in B_w with jth vertex w_1, called the *defining rectangle of $\mathscr{E}^j$ at w_1*, such that every rectangle in R_{w2} with jth vertex w_1 lies in $\mathscr{E}^j$. The j-complete families are the analogues in two dimensions of $\mathscr{L}$ and $\mathscr{R}$ for one dimension.

THEOREM 41.1. *If $\mathscr{E}^j$ is j-complete in an elementary set B_w, for $j = 1, 2, 3, 4$, then there is a division of B_w formed from the rectangles of these four families.*

Suppose the theorem false. Then we can divide B_w into a finite number of closed rectangles. If the theorem is true for each of these rectangles it is true for B_w. Thus it is false for a rectangle, say $R_{w3} = ([a, b], [u, v])$. Bisecting the rectangle by lines $x = \frac{1}{2}(a+b)$, $y = \frac{1}{2}(u+v)$, then as before, the theorem is false for at least one of the smaller rectangles. We choose the left-hand topmost one. This construction may be repeated, and we obtain a sequence

$$R_{w3} \supseteq R_{w3}^{1} \supseteq \dots \supseteq R_{w3}^{n} \supseteq \dots$$

of rectangles. Let (x_n, y_n) be the first vertex of R_{w3}^{n}. Then $\{x_n\}$ is monotone increasing and is bounded above by b, and so tends to a limit, say x. Similarly $\{y_n\}$ tends to a limit y. If R_{w4}^{j} is the defining rectangle of $\mathscr{E}^j$ at (x, y), for $j = 1, 2, 3, 4$, there is an integer n such that R_{w3}^{n} is contained in the union of the $R_{w4}^{j} (j = 1, 2, 3, 4)$. Lines through (x, y) parallel to the axes divide R_{w3}^{n} up into one, or two, or four rectangles, each lying in an $\mathscr{E}^j$ with (x, y) as the jth vertex, for an appropriate value of j. Thus the theorem is true for R_{w3}^{n}, contrary to hypothesis. Hence the theorem must be true†.

We put $\mathbf{C}_\sigma = \{\mathscr{E}_\sigma^1, \mathscr{E}_\sigma^2, \mathscr{E}_\sigma^3, \mathscr{E}_\sigma^4\}$, where σ stands for any collection of suffices; and if, in this book, we say that $\mathbf{C}_\sigma$ is complete in B_w, a two-dimensional set, we will imply that

---

† Theorem 1 shows that the collection $\mathbf{C} = \{\mathscr{E}^1, \mathscr{E}^2, \mathscr{E}^3, \mathscr{E}^4\}$ *is complete in B_w*, in the sense of Henstock (1961a) This book can be read without a knowledge of that paper.

Ⓕ

B_w is an elementary set, and that $\mathscr{E}_0{}^j$ is j-complete in $B_w (j = 1, 2, 3, 4)$.

Fubini version

Theorem 41.2. *Given* C *complete in* $\mathrm{R_w} = ([\mathrm{a, b}], [\mathrm{u, v}])$, *then to each fixed* x *in* $[\mathrm{a, b}]$ *there corresponds a division* $\mathscr{D}(\mathrm{x})$ *of* $[\mathrm{u, v}]$ *with a given specification, and there is a set* $\mathbf{A}$ *complete in* $[\mathrm{a, b}]$, *such that if* $\mathrm{I_x}$ *is in* $\mathscr{L} \cup \mathscr{R}$ *with associated point* x, *and if* $\mathrm{I_y}$ *is in* $\mathscr{D}(\mathrm{x})$ *with associated point* y, *then* $(\mathrm{I_x, I_y})$ *is in the corresponding* $\mathscr{E}^{\mathrm{j}}$ *with* j*th vertex* $(\mathrm{x, y})$.

Let $\mathscr{F}$ be a family of rectangles. If I_x is a fixed interval on the x-axis then the family of all intervals I_y on the y-axis such that (I_x, I_y) lies in $\mathscr{F}$, is called the *I_x-slice of $\mathscr{F}$*. Let x be fixed in $[a, b]$, and let (I_{x1}, I_{y1}) be the defining rectangle of $\mathscr{E}^1$ at a point (x, y). Then every interval I_{y2} in I_{y1} with top vertex y lies in the I_{x2}-slice of $\mathscr{E}^1$ for each interval I_{x2} in I_{x1} with left vertex x. It follows that the limit as $\delta \to 0$ of the $[x, x+\delta)$-slice of $\mathscr{E}^1$ is the 'left'-complete set $\mathscr{L}_1$, in an obvious notation, of intervals formed by using the I_{y1} as defining intervals. Similarly, from $\mathscr{E}^4$ we obtain a 'right'-complete family $\mathscr{R}_1$ of intervals in $[u, v]$. From $\mathscr{E}^2, \mathscr{E}^3$, using $[x-\delta, x)$-slices, we obtain, respectively $\mathscr{L}_2$ 'left'-complete and $\mathscr{R}_2$ 'right'-complete in $[u, v]$. From $\mathscr{L}_1 \cap \mathscr{L}_2$, $\mathscr{R}_1 \cap \mathscr{R}_2$ we can define a $\delta_1(x; y) > 0$ such that $[t, y)$ is in $\mathscr{L}_1 \cap \mathscr{L}_2$, $[y, t')$ in $\mathscr{R}_1 \cap \mathscr{R}_2$, for intervals of length not greater than $\delta_1(x; y)$. Suppose that to each x there corresponds a division $\mathscr{D}(x)\{I_y\}$ of $[u, v]$ such that if y is the associated point of I_y then the length of I_y is not greater than $\delta_1(x; y)$. Then to each such I_y corresponds a δ_2 such that I_y is in either the $[x, x+\delta)$-slice of $\mathscr{E}^1$ and the $[x-\delta, x)$-slice of $\mathscr{E}^2$, for all $0 < \delta \leqslant \delta_2$, when I_y is 'left'-hand; or is in the $[x-\delta, x)$-slice of $\mathscr{E}^3$ and the $[x, x+\delta)$-slice of $\mathscr{E}^4$, for all $0 < \delta \leqslant \delta_2$, when I_y is 'right'-hand. Let $\delta_3(x) > 0$ be the minimum of the δ_2 for the finite number of I_y in $\mathscr{D}(x)$. Then $\delta_3(x)$ can be used for a δ_2 independent of which of the I_y in $\mathscr{D}(x)$ is being considered, but is dependent on x, and it defines the final complete set $\mathbf{A}$ that is independent of y. Thus we have proved the theorem.

42. Functions of rectangles

A function h of rectangles R_w, or a *rectangle function*, is defined in an elementary set B_w if to each rectangle in B_w (with sides parallel to the axes) there is assigned a number. The rectangles in our divisions are all half-open, so that we can write $h(a, b; u, v)$ in place of $h([a, b), [u, v))$. To define finite additivity, superadditivity, and subadditivity, we cannot restrict ourselves to two rectangles and their union, as there is likely to occur a situation like that of Figure 1 in which the union of no two rectangles of the division gives another rectangle. Thus we use the following definitions.

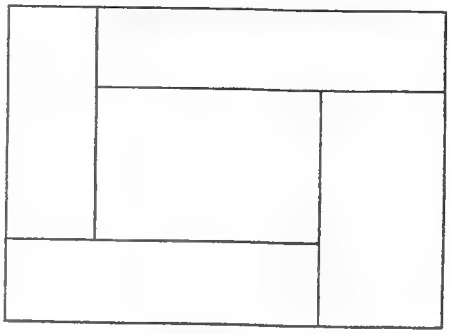

Figure 1

A rectangle function h *is finitely additive* if for each division $\mathscr{D}\{R_{w1}{}^-\}$ of every closed rectangle R_w, we have

$$(\mathscr{D}) \sum h(R_{w1}{}^-) = h(R_w{}^-) \tag{42.1}$$

If h is defined for elementary sets, too, we shall require also that if $\mathscr{D}$ is a division of an elementary set B_w, then we can always replace $R_w{}^-$ in (42.1) by $B_w{}^-$. The rectangle function h *is finitely superadditive* (respectively, *subadditive*) if in (42.1) we can always replace $=$ by $\leqslant$ (respectively, $\geqslant$).

43. The integration of rectangle functions

We now use a vector $\mathbf{h}$ whose components are four rectangle functions h_j, for $j = 1, 2, 3, 4$, defined for half-open rectangles R_w contained in an elementary set B_w. Corresponding to each division $\mathscr{D}\{R_{w1}{}^-\}$ with a given specification, so that $R_{w1}{}^-$ has its jth vertex as associated point, there is defined a sum

$$(\mathscr{D}) \sum h_j(R_{w1}{}^-) = (\mathscr{D}) \sum h_j$$

We denote by $S(\mathbf{h}; \mathbf{C}; B_w)$ the collection of such sums, for all divisions of B_w formed from the rectangles of

$$\mathscr{E}^j (j = 1, 2, 3, 4),$$

where R_{w1}^- lies in $\mathscr{E}^j$. Then the real or complex number I *is the Riemann-complete integral of* $\mathbf{h}$ *in* B_w if, given $\epsilon > 0$, we can choose $\mathbf{C}$ so that $S(\mathbf{h}; \mathbf{C}; B_w)$ lies in the closed circle with centre I and radius ϵ. The notation for I is

$$RC(\mathbf{h}; B_w) = (RC)\int_{B_w} \mathbf{h}$$

As in one dimension there are special cases. If $h_j = h\ (j = 1, 2, 3, 4)$ we replace $\mathbf{h}$ by h. If $h_j = f(x, y)h_j^*$, where (x, y) is the jth vertex of the corresponding rectangle, we put

$$\{fh_1^*, fh_2^*, fh_3^*, fh_4^*\}$$

for $\mathbf{h}$. Finally, if

$$H_g(x, x'; y, y') = g(x', y') - g(x, y') - g(x', y) + g(x, y)$$

where g is a point function, then

$$(RC)\int_{B_w} f\,\mathrm{d}g = (RC)\int_{B_w} \{fH_g, fH_g, fH_g, fH_g\}$$

where the value of f is taken at the jth vertex, for $j = 1, 2, 3, 4$, respectively.

Next, $\mathbf{h}$ *is variationally equivalent to* $\mathbf{h}^*$ *in an elementary set* B_w if, given $\epsilon > 0$, there are a $\mathbf{C}_1$ complete in B_w, and a non-negative finitely superadditive rectangle function χ_w defined also for B_w^-, such that

$$\chi_w(B_w^-) < \epsilon \tag{43.1}$$

$$|h_j(R_{w1}^-) - h_j^*(R_{w1}^-)| \leqslant \chi_w(R_{w1}^-)$$
$$(R_{w1}^- \in \mathscr{E}_1^j, j = 1, 2, 3, 4) \tag{43.2}$$

If $h_j^* = H (j = 1, 2, 3, 4)$, where H is finitely additive, then $H(B_w^-)$ is called the *variational integral of* $\mathbf{h}$ *in* B_w, written

$$(V)\int_{B_w} \mathbf{h}$$

The special cases are as for the Riemann-complete integral.

The equivalence of the two integrals follows as in Theorem 24.1 (p. 40) for one-dimensional sets.

Ex. 43.1. If g is a point function, and if $\partial^2 g/\partial x \partial y$ exists as the limit

$$\lim_{j,k\to 0} H_g(x, \underset{x}{\cancel{x}}+j; y, y+k)/jk$$

for each (x, y) in B_w, prove that the variational integral of $\partial^2 g/\partial x \partial y$ with respect to the area function in B_w, exists and is equal to $H_g(B_w^-)$.

Ex 43.2. Show that the two-dimensional analogues of the Riemann–Stieltjes and Burkill integrals of Section 3 (p. 5) are equal to the corresponding Riemann-complete integrals. Note that if in the Riemann–Stieltjes sum we use the point (ξ, η) of R_{w1}, we can draw lines through the point parallel to the x- and y-axes, dividing R_{w1} up into one, two, or four rectangles, so that each such sum is a sum using the four vertices alone as associated points.

Ex. 43.3. Prove that if See G. H. Hardy, Course of Pure Math, 4th

$$f(x, y) = 2(x-y)(x+y)^{-3}\ (x \geqslant 0, y \geqslant 0, x+y > 0),$$ Ed., 1925, p.

$$f(0, 0) = 0$$

Ex. 38. See Math Tripos, 191

then $(V)\int_0^1 \left\{(V)\int_0^1 f(x, y)\,\mathrm{d}x\right\}\mathrm{d}y = -1,$ Part II, Schedule B, Paper 2.

$$(V)\int_0^1 \left\{(V)\int_0^1 f(x, y)\,\mathrm{d}y\right\}\mathrm{d}x = 1$$

Also prove that the variational integral in $([0, 1], [0, 1])$ minus $([0, x), [0, y))$ is $(y-x)/(x+y)$, for x, y in $(0, 1)$, tending to 1 if x tends to 0 first, and then y; tending to -1 if y tends to 0 first, and then x.

Ex. 43.4. Let mI be the length of the interval I, m (*empty set*) $= 0$, and put

False

$$h(x, x'; y, y') = \begin{cases} m((x, x') \cap (y'-y, 2y'-2y))(x > 0) \\ 0 \qquad (x = 0) \end{cases}$$

If $y' - y < \frac{1}{2}x$ then $m = 0$
So need $x = 0$, & then $m = 0$ anyway.

Show by taking $y' < y + \frac{1}{2}x$ that the variational integral in $([0, 1]; [0, 1])$ is 0, whereas the variational integral in the y-direction in $[0, 1]$, of the variational integral of h in the x-direction in $[0, 1]$, is 1.

Ex. 43.5. Let $f(y)$ be a point function independent of x, and let H be a function of rectangles (I_x, I_y) that is finitely

A. L. Cauchy, 'Mémoire sur les intégrales définies', Œuvres (1) 1, 394–396.
gave $\dfrac{x^2-y^2}{(x^2+y^2)^2}$

additive in I_x. If f is variationally integrable with respect to H in $R_w = ([a, b], [u, v])$, show by using Theorem 41.2 with x and y interchanged, that

$$(V)\int_{R_w} f\,dH = (V)\int_u^v f\,dH(a, b; .)$$

Ex. 43.6. Let F be a function of a point x and an interval I_y that is finitely additive in I_y, and let $\mathbf{h}_x$ be a pair of functions of x-intervals independent of y. If

$$h_1 = h_4 = Fh_{xr},\ h_2 = h_3 = Fh_{xl}$$

and if $\mathbf{h}$ is variationally integrable in $R_w = ([a, b], [u, v])$, show by using Theorem 41.2 that the variational integral is equal to

$$(V)\int_a^b F(.; u, v)\,d\mathbf{h}_x$$

44. Fubini's theorem

In this section we give cases when a double integral is equal to a repeated integral, and when we can invert the order of integration in a repeated integral. Ex. 43.3, 43.4 show that for the second to be true we probably need the existence of the double integral, but even if it exists it is not always equal to the corresponding repeated integrals. Thus we have to restrict the rectangle functions used. For a deeper example, of a non-measurable plane set that has just one point in common with each parallel to the axes, see Sierpinski (1920).

THEOREM 44.1. *Let the pair* $\mathbf{h}_x$ *of interval functions be* VBG* *in* [a, b], *let* $\mathbf{h}_y(x; .)$ *be a pair of interval functions in* [u, v] *with parameter* x in [a, b], *and let*

$$\mathrm{h}_1(\mathrm{x}, \mathrm{x}'; \mathrm{y}, \mathrm{y}') = \mathrm{h}_{\mathrm{xr}}(\mathrm{x}, \mathrm{x}')\mathrm{h}_{\mathrm{yl}}(\mathrm{x}; \mathrm{y}, \mathrm{y}') = h_x(x, x'; x)h_y(y, y'; x, y)$$
$$\mathrm{h}_2(\mathrm{x}, \mathrm{x}'; \mathrm{y}, \mathrm{y}') = \mathrm{h}_{\mathrm{xl}}(\mathrm{x}, \mathrm{x}')\mathrm{h}_{\mathrm{yl}}(\mathrm{x}'; \mathrm{y}, \mathrm{y}')$$
$$\mathrm{h}_3(\mathrm{x}, \mathrm{x}'; \mathrm{y}, \mathrm{y}') = \mathrm{h}_{\mathrm{xl}}(\mathrm{x}, \mathrm{x}')\mathrm{h}_{\mathrm{yr}}(\mathrm{x}'; \mathrm{y}, \mathrm{y}')$$
$$\mathrm{h}_4(\mathrm{x}, \mathrm{x}'; \mathrm{y}, \mathrm{y}') = \mathrm{h}_{\mathrm{xr}}(\mathrm{x}, \mathrm{x}')\mathrm{h}_{\mathrm{yr}}(\mathrm{x}; \mathrm{y}, \mathrm{y}')$$

If $\mathbf{h}$ *has a variational integral* $\mathrm{H}(\mathrm{R_w})$ *in* $\mathrm{R_w} = ([\mathrm{a}, \mathrm{b}], [\mathrm{u}, \mathrm{v}])$, *then the point function*

$$\mathrm{J}(\mathrm{x}) = (\mathrm{V})\int_{\mathrm{u}}^{\mathrm{v}} \mathbf{h}_{\mathrm{y}}(\mathrm{x}; .) \tag{44.1}$$

T.W. Lee gives example $f(P, Q)\,h_x(P; I_x)\,h_y(P, Q), J_y$,
We can put f into h_y if we have the other form.

exists, except for x *in a set* X *with*

$$\mathrm{V}(\mathbf{h}_{\mathrm{x}}; [\mathrm{a}, \mathrm{b}]; \mathrm{X}) = 0 \tag{44.2}$$

and if we put 0 *when* J(x) *does not exist, the variational integral of*

$$\{\mathrm{J}(\mathrm{x})\mathbf{h}_{\mathrm{xl}}(\mathrm{t}, \mathrm{x}), \mathrm{J}(\mathrm{x})\mathbf{h}_{\mathrm{xr}}(\mathrm{x}, \mathrm{t}')\} \tag{44.3}$$

is $\mathrm{H}(\mathrm{R_w})$, *in* [a, b], i.e.

$$(\mathrm{V})\int_{\mathrm{R_W}} \mathbf{h}_{\mathrm{x}}\mathbf{h}_{\mathrm{y}}(\mathrm{x}; .) = (\mathrm{V})\int_{\mathrm{a}}^{\mathrm{b}} \left\{(\mathrm{V})\int_{\mathrm{u}}^{\mathrm{v}} \mathbf{h}_{\mathrm{y}}(\mathrm{x}; .)\right\} d\mathbf{h}_{\mathrm{x}}$$

Given $\epsilon > 0$, there are $\mathbf{C}$, complete in R_w, and χ_w, non-negative and finitely superadditive, such that

$$\chi_w(R_w^-) < \epsilon \tag{44.4}$$
$$|H(I_x, I_y) - h_j(I_x, I_y)| \leqslant \chi_w(I_x, I_y)$$
$$((I_x, I_y) \in \mathscr{E}^j, j = 1, 2, 3, 4) \tag{44.5}$$

As in the definition of the Riemann-complete integral, we consider the set

$$S(\mathbf{h}_y(x; .); \mathbf{A}; [u, v]) \tag{44.6}$$

of sums over divisions of $[u, v]$ from a set $\mathbf{A}$ complete in $[u, v]$. If for a given integer n and all $\mathbf{A}$ complete in $[u, v]$, (44.6) has diameter greater than $1/n$, we put x into a set X_{1n}. Then there are two sums S_1, S_2 from divisions of $[u, v]$ using $\mathbf{A}$, such that

$$|S_1 - S_2| > 1/n \tag{44.7}$$

p.102

For such an x, and the given $\mathbf{C}$, we can define the $\delta_1(x; y) > 0$ of Theorem 41.2 and so obtain an $\mathbf{A}$ complete in $[u, v]$. Let $\mathscr{D}_1$, $\mathscr{D}_2$ be two divisions of $[u, v]$ using $\mathbf{A}$ that give sums S_1, S_2 satisfying (44.7). By using one $\mathscr{D}_1$ for each x of X_{1n}, and arbitrary $\mathscr{D}$ from $\delta_1(x; y)$ for other x, we obtain in Theorem 41.2 an $\mathbf{A}_1$ complete in $[a, b]$. Similarly, using the other $\mathscr{D}_2$ for each x of X_{1n}, and arbitrary $\mathscr{D}$ from $\delta_1(x; y)$ for other x, we obtain $\mathbf{A}_2$, and then $\mathbf{A}_1$. $\mathbf{A}_2$ will be connected in the same way with both $\mathscr{D}_1$ and $\mathscr{D}_2$. Thus from (44.5), as χ_w is finitely superadditive,

$$|H(I_x; u, v) - h_{xs}(I_x).S_k| \leqslant \chi_w(I_x; u, v).$$

($k = 1, 2$; $s = l, r$ *according as* I_x *is in* $\mathscr{L}_1 \cap \mathscr{L}_2$ *or* $\mathscr{R}_1 \cap \mathscr{R}_2$),

$$|h_{xs}(I_x)|/n \leqslant |h_{xs}(I_x)|.|S_1 - S_2| \leqslant 2\chi_w(I_x; u, v)$$

This gives, using (44.4),

$$V(\mathbf{h}_x; [a, b]; X_{1n}) \leqslant 2n\chi_w(a, b; u, v) < 2n\epsilon$$

As $\epsilon > 0$ is arbitrary the variation is 0. This is true for $n = 1, 2, \ldots$, so that (44.2) follows from Theorem 28.1 (p. 49).

We now show that the variational integral of (44.3) is $H(R_w^-)$. We are given that $\mathbf{h}_x$ is VBG^*, so that by Theorem 29.1 (29.1; 29.2), p. 56, there are a non-negative finitely superadditive interval function χ_x, a point function $k(x) > 0$ and a set A_3 complete in $[a, b]$, such that

$$|h_{xs}| \leqslant k(x)\chi_x \tag{44.8}$$

for intervals of $\mathscr{L}_3 \cup \mathscr{R}_3$ with associated point x. Also, for each x not in X and each $\epsilon > 0$ there are $\delta_2(x; y) > 0$ with $\delta_2 \leqslant \delta_1$, and a non-negative finitely superadditive interval function $\chi_y(x; .)$, such that

$$\chi_y(x; u, v) < \epsilon/\{k(x)\chi_x(a, b)\} \tag{44.9}$$

$$|RC(\mathbf{h}_y(x; .); I_y) - h_{ys}(x; I_y)| \leqslant \chi_y(x; I) \tag{44.10}$$

$$(s = l,\ I_y = [y-\delta, y);\ \text{or}\ s = r,\ I_y = [y, y+\delta);$$
$$\text{with}\ 0 < \delta \leqslant \delta_2(x; y))$$

Using δ_2 there is a division $\mathscr{D}_3\{p, q\}$ of $[u, v]$ such that by (44.10),

$$|RC(\mathbf{h}_y(x; .); [u, v]) - (\mathscr{D}_3)\sum h_{ys}(x; p, q)| \leqslant \chi_y(x; u, v) \tag{44.11}$$

The integral in (44.11) is $J(x)$. We use $\mathscr{D}_3$ to define $\delta_3(x)$ for x outside X. If I_x is $[x-\delta, x)$ or $[x, x+\delta)$, with $0 < \delta \leqslant \delta_3(x)$, and if I_y is in $\mathscr{D}_3$, then (44.5) is satisfied, so that with (44.11) ① we obtain

$$|H(I_x; u, v) - J(x)h_{xs}(I_x)| \leqslant \chi_w(I_x; u, v) + \\ + \chi_y(x; u, v)|h_{xs}(I_x)| \tag{44.12}$$

where $s = l, r$ according as I_x is left-hand or right-hand.

By (44.8; 44.10; 44.12),

$$|H(I_x; u, v) - J(x)h_{xs}(I_x)| \\ \leqslant \chi_w(I_x; u, v) + \epsilon\chi_x(I_x)/\chi_x(a, b) \tag{44.13}$$

$$(s = l,\ I_x \in \mathscr{L}_3 \cap \mathscr{L}_4,\ \text{or}\ s = r,\ I_x \in \mathscr{R}_3 \cap \mathscr{R}_4)$$

where A_4 is formed from $\delta_3(x)$. The sum over a division of $[a, b]$ of the right-hand side of (44.13) is bounded by 2ϵ, so that the variational integral of (44.3) is $H(R_w)$.

There are two important groups which are special cases of Theorem 44.1. First we have those connected with Fubini's theorem. This could be a helpful special case

THEOREM 44.2. *Let* φ_x, φ_y *be* VBG* *interval functions in* [a, b], [u, v], *respectively, and let* f(x, y) *be a point function in* $R_w = ([a, b]; [u, v])$. *If*

$$h_j(I_x, I_y) = f(x, y)\varphi_x(I_x)\varphi_y(I_y)$$

where (x, y) *is the* j*th vertex of* (I_x, I_y) *and if* **h** *has a variational integral* $H(R_w^-)$ *in* R_w, *then the point functions*

$$J(x) = (V)\int_u^v f(x, .)\, d\varphi_y,\quad K(y) = (V)\int_a^b f(., y)\, d\varphi_x$$

exist, except for x *in* X, y *in* Y, *with*

$$V(\varphi_x; [a, b]; X) = 0 = V(\varphi_y; [u, v]; Y) \tag{44.14}$$

and also

$$(V)\int_a^b J\, d\varphi_x = (V)\int_u^v K\, d\varphi_y = H(R_w) = (V)\int_{R_w} f\, d\varphi_x\varphi_y$$

If also **h** *is* VB* *in* R_w, *or if* φ_x, φ_y *are* VB* *and*

$$H^+(R_w^-) = (V)\int_{R_w} |f|\, dV(\varphi_x; .)V(\varphi_y; .)$$

exists, then

$$J^+(x) = (V)\int_u^v |f|\, dV(\varphi_y; .),\quad K^+(y) = (V)\int_a^b |f|\, dV(\varphi_x; .)$$

exist, except for x *in* X, y *in* Y, *satisfying* (44.14), *and*

$$(V)\int_a^b J^+\, dV(\varphi_x; .) = (V)\int_u^v K^+\, dV(\varphi_y; .) = H^+(R_w)$$

In particular, φ_x *and* φ_y *can be read as point functions. More generally, they can be read as pairs of interval functions.*
This includes the example at the bottom of p. 367 of W. H. Young (1910).

Ex 17.3 **THEOREM 44.3.** *Let* $\mathbf{h}_x, \mathbf{h}_y$ *be* VB* *in* $[a, b], [u, v]$, *respectively, and let* W_w *be a plane set. Denote by* $W_y(x)$, $W_x(y)$ *the sets, respectively, of* y *for fixed* x, *and of* x *for fixed* y, *such that* (x, y) *is in* W_w. *In Theorem* 44.1 *let* $\mathbf{h}_1, \mathbf{h}_2$ *be the vectors obtained by replacing* $h_{ys}(x; .)$ *respectively by*

$$\mathrm{ch}(W_w; (x, \eta))h_{ys},\ h_{ys}$$

where η *is the associated point of the* y-*interval. If* $|\mathbf{h}_1|$ *is variationally integrable in* $R_w = ([a, b]; [u, v])$, *then*

$$V(\mathbf{h}_2; R_w; W_w) = (V)\int_a^b V(\mathbf{h}_y; [u, v]; W_y(.))\, d\mathbf{h}_x$$

$$= (V)\int_u^v V(\mathbf{h}_x; [a, b]; W_x(.))\, d\mathbf{h}_y$$

In Theorem 44.3 we use Theorems 44.1 and 31.2 (31.9), p.61 and the plane analogue of the latter.

The second group of theorems are connected with some due to Cameron and Martin (1941), and Robbins (1948).

Helpful special case.

THEOREM 44.4. *Let* f *be a point function in* $[u, v]$, *and let* $F(x; I_y)$, $G(I_x)$ *be interval functions in* $[u, v], [a, b]$, *respectively, the first interval function containing a parameter* x *that lies in* $[a, b]$, *and the second being* VBG*. *If*

$$h_j(I_x, I_y) = G(I_x)f(y)\,F(x; I_y)$$

where (x, y) *is the* j*th vertex of* (I_x, I_y), *and if* $\mathbf{h}$ *is variationally integrable in* $R_w = ([a, b]; [u, v])$ *to* $H(R_w^-)$, *then*

$$J(x) = (V)\int_u^v f\, dF(x; .)$$

exists, except for a set X *of* x *satisfying* (44.2), *and*

$$(V)\int_a^b J\, dG = (V)\int_a^b \left\{(V)\int_u^v f\, dF(x; .)\right\} dG$$

exists and is equal to $H(R_w^-)$. *If also* $|\mathbf{h}|$ *is variationally integrable in* R_w *to* $H^+(R_w^-)$ *then*

$$J^+(x) = (V)\int_u^v |f|\, dV(F(x; .); .)$$

exists except for a set X *of* x *satisfying* (44.2), *and*

$$(V)\int_a^b J^+\, dV(G; .)$$

exists and is equal to $H^+(R_w^-)$.

Helpful special case.

THEOREM 44.5. *If in Theorem* 44.4, $F(x; I_y)$ *is finitely additive in* I_y, *and if*

$$(V)\int_a^b F(.\,; I_y)\, dG$$

exists for each I_y *in* $[u, v]$, *then the integrability of* $\mathbf{h}$ *in* R_w *implies the existence of*

$$(V)\int_u^v f\, d\left\{(V)\int_a^b F(.\,; I_y)\, dG\right\} = H(R_w^-)$$

and similarly for the absolute form.
Theorem 31.3 (31.20), p. 63 (integration by substitution) for two dimensions gives

$$(V)\int_{R_w} \mathbf{h} = (V)\int_{R_w} f(y)\, \mathrm{d}(V)\int_{R'_w} F\, \mathrm{d}G$$

where R_w' is the rectangle of integration in R_w. As $f(y)$ is independent of x, as the variational integral is finitely additive and as F is finitely additive in I_y, we obtain by Ex. 43.5; 43.6, pp.105-6,

$$(V)\int_{R_w} \mathbf{h} = (V)\int_u^v f(y)\, \mathrm{d}\left\{(V)\int_{([a,b];I_y)} F\, \mathrm{d}G\right\}$$

$$= (V)\int_u^v f\, \mathrm{d}\left\{(V)\int_a^b F(.\,; I_y)\, \mathrm{d}G\right\}$$

Theorems 44.4 and 44.5 together give the 'Unsymmetrical Fubini Theorem'. The integrals of Cameron et al (1941) and Robbins (1948) are over an infinite range, but this will make no difference to the theorems, as will be pointed out in Chapter 7, and the discussion will be completed there.

*45. Further results

Theorems of previous chapters can be generalized to two dimensions, apart from 22.2, and those of Sections 23, 32 and 33. But Sierpinski's Lemma is replaced by Vitali's theorem, and we cannot prove that a two-dimensional set of inner variation zero is of variation zero. Thus when we differentiate the integral

$$(V)\int f\,d\mathbf{h} \tag{45.1}$$

the exceptional set of inner variation 0 is not always of variation 0, and in the case when $h_1, \ldots, h_4$ are all the area function we have to assume the integrability of $|f|\max(\log|f|; 0)$. For Vitali's theorem see, for example, Saks (1937, pp. 109–112) and Jeffery (1951, pp. 110–114). For the strong differentiation theorem of Jessen, Marcinkiewicz and Zygmund see, for example, Saks, 1937, pp. 147–9).

The two-dimensional integration by substitution is interesting, in that it is enough to assume in (45.1) that

$$h_j = (V)\int g\,d\mathbf{h}^*\ (j = 1, 2, 3, 4) \tag{45.2}$$

obtaining by the analogue of Theorem 31.3 (31.20), p. 63,

$$(V)\int fg\,d\mathbf{h}^* \tag{45.3}$$

In the special case when h_j is the area function in the (x, y) plane, and h_j^* that in the (u, v) plane for $j = 1, 2, 3, 4$, where $x = x(u, v)$, $y = y(u, v)$ are suitable functions, then g is the two-dimensional derivative $D(\mathbf{h}, \mathbf{h}^*; B; (u, v))$, which by a well-known proof is the modulus of the Jacobian

$$\partial(x, y)/\partial(u, v)$$

Lack of space forbids further elaboration.

7

THE CAUCHY AND DENJOY EXTENSIONS, AND THE INTEGRAL IN AN INFINITE RANGE

46. The Cauchy extension

On previous pages all our integrals have been in a finite interval or a finite rectangle. But in practice we have to integrate in many other sets. We have already integrated in a set X, each point of which lies inside an interval I, simply by introducing the characteristic function of X in the appropriate place and integrating in I. The next step is to integrate in an interval when difficulties occur at one or both ends of the interval. When the integral $H(a, u)$ exists for each u in $a < u < b$, then A. L. Cauchy (1789–1857) defined $H(a, b)$ to be the limit of $H(a, u)$ as $u \to b$, if the limit exists. We call such an extension a *Cauchy extension* of the original integral. It is important to note that if we take into account any discontinuity of the interval functions at b, then the Riemann-complete integral already contains such a Cauchy extension, as Theorem 46.1 shows.

THEOREM 46.1. *If* $\mathrm{H(u, v) = RC(\mathbf{h}; [u, v])}$ *exists and is finite for each* u, v *in* $\mathrm{a \leqslant u < v < b}$, *and if, for each* u *in* $\mathrm{a \leqslant u < b}$,

$$\mathrm{H_1(u, b)} = \lim_{\epsilon\to 0+}\{\mathrm{H(u, b-\epsilon) + h_1(b-\epsilon, b)}\} \tag{46.1}$$

exists, then $\mathrm{H(a, b)}$ *exists and is equal to* $\mathrm{H_1(a, b)}$.

Thus (46.1) is sufficient for the existence of $\mathrm{H(a, b)}$ in this case. By Theorem 21.2 (21.12; 21.13), p. 33, it is also necessary.

Putting $u_j = b-(b-a)2^{-j}(j = 0, 1, 2, \ldots)$, we are given that $H(u_{j-1}, u_j)$ exists and is finite for $j = 1, 2, \ldots$. Given $\epsilon > 0$, then there are $\mathbf{A}_{1j}$ complete in, and formed from intervals in $[u_{j-1}, u_j]$, such that

$$S(\mathbf{h}; \mathbf{A}_{1j}; [u_{j-1}, u_j]) \tag{46.2}$$

lies in the closed circle centre $H(u_{j-1}, u_j)$ and radius $\epsilon.2^{-j-1}$.

Also, by definition of $H_1(u, b)$ there is a $\delta > 0$ such that if $b-\delta \leqslant v < b$ then

$$|H(u, v)+h_l(v, b)-H_1(u, b)| < \tfrac{1}{2}\epsilon$$

and this, by the finite additivity of H gives

$$|H_1(v, b)-h_l(v, b)| < \tfrac{1}{2}\epsilon \tag{46.3}$$

Let $\mathscr{L}$ be the union of the $\mathscr{L}_{1j}$, together with intervals $[v, b)$ for $b-\delta \leqslant v < b$, and let $\mathscr{R}$ be the union of the $\mathscr{R}_{1j}$. Then **A** is complete in $[a, b]$, so that we can obtain divisions $\mathscr{D}$ of $[a, b]$ from **A**. By construction such a division can be split up into divisions $\mathscr{D}_{1j}$ of $[u_{j-1}, u_j]$ $(j = 1, 2, \ldots, J-1)$, for some integer J, together with an interval $[v, b)$ with $u_{J-1} < v \leqslant u_J$, and a division $\mathscr{D}_{1J}$ of $[u_{J-1}, v]$. Using (46.3), the results for sets (46.2), and the proof of Theorem 21.2 (p. 32),

$$(\mathscr{D})\sum h_s = \sum_{j=1}^{J} (\mathscr{D}_{1j})\sum h_s+h_l(v, b)$$

$$\left|\sum_{j=1}^{J} (\mathscr{D}_{1j})\sum (h_s-H)+\{h_l(v, b)-H_1(v, b)\}\right|$$

{not slanting
{bars

$$< \sum_{j=1}^{J-1} \epsilon\,.\,2^{-j-1}+\epsilon\,.\,2^{-J}+\epsilon\,.\,2^{-1} = \epsilon$$

By the finite additivity of H we see that $H(a, b)$ exists and is equal to

$$H(a, v)+H_1(v, b)$$

this sum being independent of v in $a < v < b$.

Note that if we use the characteristic function of $[a, u]$, integrate in $[a, b]$, and then take the limit, the result is the integral in $[a, b)$, if it exists, and not the integral in $[a, b]$. The point b adds the limit (46.1).

Similar results occur if the Riemann-complete integral exists in $[u, b]$, for each u in $a < u < b$, and if the limit corresponding to (46.1) exists; or if the Riemann-complete integral exists in $[u, v]$, for each u, v in $a < u < v < b$, with two appropriate limits occurring as u, v tend independently and respectively to a, b; or if we have similar situations in two dimensions.

a+, b- →

a+ b- →

The situation is different when $u \to a$, $v \to b$ simultaneously, for example, when $u-a = b-v$. This can occur in integration obtained by using complex contours and the calculus of residues, and the result is a genuine extension of the Riemann-complete integral, as Ex. 46.1 shows. Thus not all the theorems need be true for this extension.

One or more conditions for the existence of

$$\lim_{u\to b-} H(a, u)$$

are usually given in textbooks on elementary analysis, and these conditions are relevant here to tell us when some Riemann-complete integrals exist.

Ex. 46.1. Let **h** be Riemann-complete integrable in $[0, u]$ for each u in $0 < u < 1$, but not for $u = 1$, and put

$$h_l(u, 1) = 0 \quad (0 \leqslant u < 1)$$

If we define

$$h_l(-v, -u) = -h_r(u, v),$$

$$h_r(-v, -u) = -h_l(u, v) \quad (0 \leqslant u < v \leqslant 1)$$

show that

$$\lim_{u\to 1} RC\,(\mathbf{h};[-u, u]) = 0$$

and give a concrete example satisfying the first part of the exercise.

47. The integral in an infinite range

It is usual to define the integral in the range $[a, +\infty)$ from a to $+\infty$ as the limit as $b \to +\infty$, of the integral in $[a, b]$, if the limit exists. This is clearly an extension of the Riemann-complete integral as we have defined it, but by a simple transformation we bring the integral into the area covered by the previous chapters. The function

$$t(x) = x/(1-x^2) \quad (-1 < x < +1)$$

is 0 at $x = 0$, tends to $+\infty$ as x tends to 1 from below, and to $-\infty$ as x tends to -1 from above. Thus we can put u', v', a', b' as the numbers with

$$t(u') = u,\ t(v') = v,\ \ t(a') = a,\ t(b') = b,$$

and then define

$$h_s{}^*(u', v') = h_s(u, v)\ (s = l, r),\ \ h_l{}^*(u', 1) = 0$$

By definition of the Riemann-complete integrals,

$$RC(\mathbf{h}^*;[a',b']) = RC(\mathbf{h};[a,b]) \qquad (47.1)$$

the first integral exists as we are supposing that the second does. By definition, (47.1), and Theorem 46.1,

$$RC(\mathbf{h};[a,+\infty)) = \lim_{b\to\infty} RC(\mathbf{h};[a,b])$$

$$= \lim_{b'\to 1} RC(\mathbf{h}^*;[a',b']) = RC(\mathbf{h}^*;[a',1]) \qquad (47.2)$$

Similarly we can define

$$RC(\mathbf{h};(-\infty,b]) = \lim_{a\to-\infty} RC(\mathbf{h};[a,b]),$$

$$RC(\mathbf{h};(-\infty,+\infty)) = \lim_{a\to-\infty}\lim_{b\to+\infty} RC(\mathbf{h};[a,b])$$

and if we suppose that $h_r^*(-1,v) = 0 (-1 < v \leqslant 0)$ we obtain

$$RC(\mathbf{h};(-\infty,b]) = RC(\mathbf{h}^*;[-1,b']),$$

$$RC(\mathbf{h};(-\infty,+\infty)) = RC(\mathbf{h}^*;[-1,+1]) \qquad (47.3)$$

Similar results hold in two dimensions. In this way we can show that all theorems using a finite range hold also for an infinite range, without having to take limits as

$$a \to -\infty,\ b \to +\infty$$

In particular we can say that **A** *is complete in* $[a,+\infty)$, if **A** is complete in $[a,b]$ for each $b > a$, and if $\mathscr{L}$ contains infinite intervals $[v,+\infty)$ for all v not less than some fixed number. These remarks are particularly important where otherwise we would have to consider iterated limits. For example, *all theorems of Chapter* 5 *are still true if we replace* E *by* $[a,+\infty)$, $(-\infty,b]$, $(-\infty,+\infty)$, *except that the step functions of Section* 39 (pp. 92–7) *are constant in some interval* $[v,+\infty)$ *and in some interval* $(-\infty,u]$. Similarly for all theorems of Section 44 (pp. 106–11), so that in particular we obtain a theorem that includes the theorem of Cameron, Martin and Robbins.

THEOREM 47.1.

$$(\mathrm{V})\int_{-\infty}^{\infty}\left\{(\mathrm{V})\int_{-\infty}^{\infty}\mathrm{f(y)}\,d_{\mathrm{y}}\mathrm{F(x;y)}\right\}d_{\mathrm{x}}\mathrm{G(x)}$$

$$= (\mathrm{V})\int_{-\infty}^{\infty}\mathrm{f(y)}\,d_{\mathrm{y}}\left\{(\mathrm{V})\int_{-\infty}^{\infty}\mathrm{F(x;y)}\,d_{\mathrm{x}}\mathrm{G(x)}\right\}$$

when the corresponding double integral exists and the inner integral on the right-hand side exists.

Ex. 47.1. Prove the following successive results.

$$x^a \leqslant n!2^a \exp(\tfrac{1}{2}x)\ (x \geqslant 2,\ a \leqslant n)$$

$x^a y^b e^{-x-y}$ is variationally integrable in $x \geqslant 1, y \geqslant 1$; if $a > -1, b > -1$ the result is true for $x \geqslant 0, y \geqslant 0$; for the transformation

$$x = uv,\ y = u(1-v)\ (u \geqslant 0,\ 0 \leqslant v \leqslant 1)$$

the modulus of the Jacobian is

$$\left|\frac{\partial(x,y)}{\partial(u,v)}\right| = u;$$

$$(V)\int_0^\infty x^a e^{-x}\,dx\,.\,(V)\int_0^\infty y^b e^{-y}\,dy$$

$$= (V)\int_0^\infty u^{a+b-1}e^{-u}\,du\,.\,(V)\int_0^1 v^a(1-v)^b\,dv$$

$$(a > -1,\ b > -1)$$

Note that the results can be proved directly, just as if $[0,\infty)$ were a finite interval. The final result is the well known functional relation

$$\Gamma(a+1)\Gamma(b+1) = \Gamma(a+b+2)B(a+1,b+1)$$

$$(a > -1,\ b > -1) \qquad (47.4)$$

between the beta and gamma functions, usually written with a, b for $a+1, b+1$ respectively, from which we derive

$$\Gamma^2(\tfrac{1}{2}) = \Gamma(1)B(\tfrac{1}{2},\tfrac{1}{2}) = B(\tfrac{1}{2},\tfrac{1}{2})$$

The transformations $v = \sin^2\theta$, $x = \frac{1}{2}t^2$, and the non-negativeness of the integrand give

$$e^{-\frac{1}{2}t^2}$$

$$\Gamma(\tfrac{1}{2}) = \pi^{\frac{1}{2}},\ (V)\int_{-\infty}^{\infty} e^{-\frac{1}{2}t^2}\,\mathrm{d}t = (2\pi)^{-\frac{1}{2}} \tag{47.5}$$

Ex. 47.2. By using Ex. 33.1 (p. 73) for finite intervals, and Theorem 28.1 (p. 49) show that if f and g have finite derivatives in $[a, \infty)$ then

$$V(H_f H_g; [a, \infty)) = 0$$

By the transformation $x = t(x')$ this corresponds to a variation in $[a', 1)$ with respect to $[a', 1]$. Hence or otherwise show that if also

$$V(H_f H_g; [a, \infty]) = 0$$

then there are numbers f_1, g_1, such that

$$\{f(x)-f_1\}\{g(x)-g_1\} \to 0 \text{ as } x \to \infty,$$

Ex. 47.3. If one of the integrals in

$$(V)\int_a^{\infty} f\,\mathrm{d}g + (V)\int_a^{\infty} g\,\mathrm{d}f = \cancel{f_1 g_1} - f(a)g(a)$$

exists then so does the other and we have the equality.

Ex 47.4. Prove by integrating by parts in $[0, \infty]$, that if $a > 0$,

$$\Gamma(a+1) = (V)\int_0^{\infty} x^a e^{-x}\,\mathrm{d}x = a\Gamma(a) \tag{47.6}$$

⋆ 48. The Denjoy extension

The Denjoy extension is given in the following theorem, in which it is shown that the extension does not give us anything more general than what we already have.

THEOREM 48.1. *Let* $\mathbf{h}$ *be a pair of interval functions and let* F, *containing* a *and* b, *be a closed set in* [a, b], *with*

$$\mathrm{G} = (\mathrm{a}, \mathrm{b}) \cap \mathrm{CF}$$

which by Theorems 15.1 (p. 20); 14.2; 14.1 (p. 19), *is a union of disjoint intervals, say,*

$$(\mathrm{u}_1, \mathrm{v}_1), \ldots, (\mathrm{u}_j, \mathrm{v}_j), \ldots$$

If there exist

$$\mathrm{RC}(\{\mathrm{ch}(\mathrm{F})\mathrm{h}_l, \mathrm{ch}(\mathrm{F})\mathrm{h}_r\}; [\mathrm{a}, \mathrm{b}]) \tag{48.1}$$

$$\mathrm{RC}(\{\mathrm{ch}(\mathrm{G})\mathrm{h}_l, \mathrm{ch}(\mathrm{G})\mathrm{h}_r\}; [\mathrm{u}_j, \mathrm{v}_j])\ (\mathrm{j} = 1, 2, 3, \ldots) \tag{48.2}$$

and if, given $\epsilon > 0$, *there is an integer* J *such that for every finite collection* $\mathscr{Q}$ *of intervals* [u, v], *each contained in* $[\mathrm{u}_j, \mathrm{v}_j]$ *for some* $\mathrm{j} \geqslant \mathrm{J}$, *no two intervals* [u, v] *being in the same* $[\mathrm{u}_j, \mathrm{v}_j]$, *we have*

$$|(\mathscr{Q})\sum \mathrm{RC}(\{\mathrm{ch}(\mathrm{G})\mathrm{h}_l, \mathrm{ch}(\mathrm{G})\mathrm{h}_r\}; [\mathrm{u}, \mathrm{v}])| < \epsilon \tag{48.3}$$

then there exists

$$\mathrm{RC}(\mathbf{h}; [\mathrm{a}, \mathrm{b}]) = \mathrm{RC}(\{\mathrm{ch}(\mathrm{F})\mathrm{h}_l, \mathrm{ch}(\mathrm{F})\mathrm{h}_r\}; [\mathrm{a}, \mathrm{b}]) + \sum_{j=1}^{\infty} \mathrm{RC}(\{\mathrm{ch}(\mathrm{G})\mathrm{h}_l, \mathrm{ch}(\mathrm{G})\mathrm{h}_r\}; [\mathrm{u}_j, \mathrm{v}_j]) \tag{48.4}$$

Note that (48.1) is the integral of $\mathbf{h}$ in F, with respect to $[a, b]$. Also, the intervals in which we integrate are closed, so that to integrate in an open interval we use the characteristic function of the open interval; the omission of a single point may alter the integral if $\mathbf{h}$ is not continuous there.

As the Riemann-complete integral is a distributive functional we can subtract (48.1) from both sides of (48.4), so that we need only prove that from (48.2; 48.3) follow the existence and equality of

$$RC(\{ch(G)h_l, ch(G)h_r\}; [a, b]) = H_1(a, b) \tag{48.5}$$

where for each elementary set E,

$$H_1(E) = \sum_{j=1}^{\infty} RC(\{ch(G)h_l, ch(G)h_r\}; [u_j, v_j] \cap E) \tag{48.6}$$

As E is an elementary set, $[u_j, v_j] \cap E$ is either empty or $[u_j, v_j]$, for all but a finite number of j. Thus from (48.3) the sequence of partial sums of the infinite series on the right of (48.6) is fundamental and so convergent, so that H_1 exists. Further, there are $\mathbf{A}_{1j}$, complete in $[u_j, v_j]$, such that the

$$S(\{ch(G)h_l, ch(G)h_r\}; \mathbf{A}_{1j}; [u_j, v_j]) \tag{48.7}$$

lie in the circles with centres $H_1(u_j, v_j)$ and radii $\epsilon.2^{-j}$, $j = 1, 2, \ldots$ so that from the proof of Theorem 21.2, (p. 32),

$$\text{the } P(\{ch(G)h_l - H_1, ch(G)h_r - H_1\}; \mathbf{A}_{1j}; [u_j, v_j]) \tag{48.8}$$

lie in the circles with centre the origin and radii $\epsilon.2^{-j+1}$.

I

From the $\mathbf{A}_{1j}$ we construct $\mathbf{A}$, complete in $[a, b]$, in the following way. The intervals of $\mathscr{L}$ with right-hand end-point x in $(u_j, v_j]$ are those intervals of $\mathscr{L}_{1j}$ with right-hand end-point x, that lie in $(u_j, v_j]$, for $j = 1, 2, \ldots$. If x is in F and is not a v_j, we include in $\mathscr{L}$ all $[t, x)$ that lie in $[a, b]$ but do not overlap with (u_j, v_j) for $j < J$. The construction of $\mathscr{R}$ is obtained by obvious changes. Let $\mathscr{D}$ be a division of $[a, b]$ from $\mathbf{A}$, and let $[t, x)$ be in $\mathscr{D}$ and $\mathscr{L}$. If x is in F and is not a v_j then

$$ch(G; x)h_1(t, x) = 0$$

and $[t, x)$ can only overlap with (u_j, v_j) if $j \geqslant J$. Thus the sum of $H_1(t, x)$, for such $[t, x)$, together with the $H_1(x, w)$ for $[x, w)$ in $\mathscr{D}$ and $\mathscr{R}$, with x in F and not a u_j, will be the limit of a sequence of sums over $\mathscr{Q}$ satisfying (48.3), and the modulus will be not greater than ϵ. If, on the other hand, $[t, x)$ lies in $\mathscr{D}$ and $\mathscr{L}$, with x at a v_j or in G, then for some integer j, x is in $(u_j, v_j]$, and $[t, x)$ is in $\mathscr{L}_{1j}$. Similarly, if $[x, w)$ lies in $\mathscr{D}$ and $\mathscr{R}$, with x at a u_j, or in G. By the construction of $\mathbf{A}$, we see that

$$[u_1, v_1), \ldots, [u_{J-1}, v_{J-1})$$

are covered wholly and

$$[u_J, v_J), [u_{J+1}, v_{J+1}), \ldots$$

partially, or wholly by these $[t, x)$, $[x, w)$. Thus by (48.6), the results for sets (48.7; 48.8), and the previous paragraph,

$$\begin{aligned}&|(\mathscr{D}) \sum ch(G)h_s - H_1(a, b)| \\ &\quad = |(\mathscr{D}) \sum \{ch(G)h_s - H_1\}| \leqslant \sum_{j=1}^{J-1} \epsilon \,.\, 2^{-j} + \sum_{j=J}^{\infty} \epsilon \,.\, 2^{-j+1} + \epsilon < 3\epsilon\end{aligned}$$

which proves the theorem. A similar result holds in higher dimensions.

8*

CONNECTIONS WITH EARLIER INTEGRALS

49. The integrals of Sections 2 to 7

The following integrals have already been proved to be included in the one-dimensional Riemann-complete integral:

the Calculus integral (Ex. 17.2, p. 25) (49.1)

the Riemann and Riemann–Stieltjes integrals (Ex. 17.3; 17.4) p.25 (49.2)

the Pollard–Getchell integral (Ex. 17.5) p.25 (49.3)

the Burkill integral of interval functions (Ex. 18.1) p.26 (49.4)

We now show that the Lebesgue and special Denjoy, and the Radon and special Denjoy–Stieltjes integrals (for a continuous monotone increasing integrator) are also included. To prove this we may naturally assume any relevant theory of the older integrals that is required.

THEOREM 49.1. *Let* h *be a continuous monotone increasing point function, let* E *be an elementary set, and let* G *be an open set. Then* G ∩ E *is a union of disjoint intervals*

$$(u_j, v_j) \quad (j = 1, 2, \ldots)$$

together possibly with a finite number of points, and

$$V(h; E; G) = \sum_{j=1}^{\infty} \{h(v_j) - h(u_j)\} \tag{49.5}$$

the h*-measure of* G ∩ E. *Further, if* X *is a set of* x, *and* h* *the outer* h*-measure,*

$$V(h; E; X) \leqslant h^*(X \cap E) \tag{49.6}$$

To prove (49.5) we use Theorems 14.1 (p. 19), 28.3 (p. 51); 31.1 (31.3), p. 60; and Ex. 28.2 (p. 55). To prove (49.6) we need only use Ex. 26.6 (p. 48) and the definition of $h^*(X \cap E)$ as the lower bound of the h-measure of G, for all open sets $G \supseteq X \cap E$, and (49.5).

Theorem 49.2. *Every Lebesgue or Radon integral of a point function with respect to a continuous point function integrator, is also a Riemann-complete integral.*

If f, g are point functions, with f bounded, g continuous and of bounded variation, and f measurable with respect to the variation of g, then the Radon integral of f with respect to g in an interval I is defined by

$$(\textit{Radon})\int_I f\,\mathrm{d}g = (\textit{Radon})\int_I f\,\mathrm{d}g_1 - (\textit{Radon})\int_I f\,\mathrm{d}g_2$$

where g_1, g_2 are monotone increasing, with $g_1 - g_2 = g$. Thus we can assume that g is continuous and monotone increasing, in which case there is a bounded sequence $\{f_n\}$ of step functions convergent to f, except possibly in a set of outer g-measure zero. By Theorem 49.1 (49.6) this set is also of variation zero, with respect to g and I. The Radon integral of each step function with respect to the continuous and monotone increasing g is equal to the corresponding Riemann–Stieltjes integral, which in turn is equal, by (49.2), to the corresponding Riemann-complete integral. Theorem 37.1 (p. 85) completes the proof when f is bounded.

When f is unbounded but non-negative, and g is continuous and monotone increasing, then $\min(f, n)$ is bounded and g-measurable, and so Radon integrable, with respect to g. By the first part the corresponding Riemann-complete integral exists with the same value. As $n \to \infty$, the value tends to the Riemann-complete integral of f with respect to g in I, by Theorem 36.1 (p. 82), and to the corresponding Radon integral by the earlier theory. Finally, for the general f, g, we split f up into its positive and negative parts, and g up as the difference of two monotone increasing functions, so that the theorem is proved. Note that the Lebesgue integral is the special case when g is x itself, and that to integrate over a set X we need only multiply f by the characteristic function of X.

Theorem 49.3. *The special Denjoy and Denjoy–Stieltjes integrals for a continuous integrator of bounded variation, are all included in the Riemann-complete integral.*

The Cauchy and Denjoy extensions, given in Theorems 46.1 (p. 113); 48.1 (p. 118), are precisely the extensions Denjoy used to define his special Denjoy integral from the Lebesgue integral, and similarly for the Stieltjes form. Hence the theorem.

The general Denjoy integral uses a different extension, and is not included in the Riemann-complete integral.

50. The Perron and Ward integrals

Include.

Perron's integral is defined as follows. Given a finite point function f, a continuous point function h *is a Perron major function of f in $[a, b]$*, if

$$\underline{D}\,h \geqslant f \tag{50.1}$$

where $\underline{D}$ denotes the lower derivative. Also h *is a Perron minor function of f in $[a, b]$*, if $-h$ is a Perron major function of $-f$ in $[a, b]$. Then

$$\overline{H}(a, b) = \inf\{h(b)-h(a)\}$$

for all Perron major functions of f in $[a, b]$, is the *upper Perron integral* while

$$\underline{H}(a, b) = \sup\{h(b)-h(a)\}$$

for all Perron minor functions of f in $[a, b]$, is the *lower Perron integral*. If the upper and lower Perron integrals are equal, they are the *Perron integral*.

Ward's definition of his integral is as follows. Let f, g be finite point functions in $a \leqslant x \leqslant b$. A point function h *is a Ward major function of f, g in $[a, b]$* if $h(a) = 0$, and if for $a \leqslant x \leqslant b$ there is a $\delta(x) > 0$ such that

$$h(u) \geqslant h(x)+f(x)\{g(u)-g(x)\} \qquad (0 \leqslant u-x \leqslant \delta(x),\ a \leqslant u \leqslant b) \tag{50.2}$$

$$h(t) \leqslant h(x)+f(x)\{g(t)-g(x)\} \qquad (0 \leqslant x-t \leqslant \delta(x),\ a \leqslant t \leqslant b) \tag{50.3}$$

The point function h *is a Ward minor function of f, g in $[a, b]$* if $-h$ is a Ward major function of $-f$, g. *The upper and lower Ward integrals* are

$$\inf h(b),\ \sup h_1(b)$$

for all Ward major functions h and all Ward minor functions h_1, of f, g in $[a, b]$. If the upper and lower Ward integrals are equal their common value is the *Ward integral of f, g in* $[a, b]$.

THEOREM 50.1. *If when* g *is* x *itself we restrict the Ward major and minor functions to be continuous, then the upper and lower Ward integrals of* f, x *are respectively equal to the upper and lower Perron integrals of* f.

The restriction to continuous functions will later be shown to be unnecessary. If h satisfies (50.2; 50.3) when g is x then h satisfies (50.1). Conversely, let $\epsilon > 0$. Then from (50.1),

$$\underline{D}h > f-\epsilon$$

so that $h(x)+\epsilon x$ satisfies (50.2; 50.3) for some $\delta(x) > 0$, and

$$h(x)-h(a)+\epsilon(x-a)$$

is a Ward major function of f, x in $[a, b]$. As $\epsilon(b-a)$ is as small as we please, we have proved that the upper Ward integral of f, x is the same as the upper Perron integral of f. Similarly for the lower integrals.

Note that in the definition of the Ward integral we have again used left and right intervals. In Saks (1937, p. 207), the small intervals are only supposed to contain x, so that in this case the associated point of an interval can lie anywhere in the interval. If we generalized this, we would not have just two interval functions but an infinity; that is, an interval function depending on a parameter that can take an infinity of values, and the theory of these, though following that given in previous chapters, would be more complicated to describe.

After writing the bulk of this book I received a letter from Professor T. H. Hildebrandt, who very kindly gave me the references Hildebrandt (1926) and Young (1915), which in turn gave the reference N. Lusin (1911–12), in which Lusin gave a proof of my Theorem 16.1 (p. 22), so that he has the priority. However, no Denjoy-depth integration is based on that theorem, so that Henstock (1960a–1961b) appears to have the priority for the integrals.

The next theorem shows that for point functions the Ward and variational integrals are equivalent.

THEOREM 50.2. *If for* $a \leqslant x \leqslant b$ *there exists*

$$(Ward)\int_a^x f\,dg = H(x) \tag{50.4}$$

with H(a) = 0, *then* H *is variationally equivalent to* f, g *in* [a, b]. *Conversely, if* H(a) = 0, *and if* H *is variationally equivalent to* f, g *when we use finitely additive* χ, *then* H *is the Ward integral of* f, g *in* [a, b].

If (50.4) exists for $a < x \leqslant b$, then, given $\epsilon > 0$, there are h, h', respectively a Ward major and a Ward minor function of f, g, such that

$$h(b)-h'(b) < \epsilon \tag{50.5}$$

Let $\delta_1(x) > 0$ be the smaller of the $\delta(x)$ for h, h' at each point x. Then if $0 < u-x \leqslant \delta_1(x)$, $0 < x-t \leqslant \delta_1(x)$, with u, t in $[a, b]$,

$$\begin{aligned}&\{h(u)-h'(u)\}-\{h(x)-h'(x)\}\\ &=[h(u)-h(x)-f(x)\{g(u)-g(x)\}]-\\ &\qquad[h'(u)-h'(x)-f(x)\{g(u)-g(x)\}] \geqslant 0,\\ &\qquad\{h(x)-h'(x)\}-\{h(t)-h'(t)\} \geqslant 0\end{aligned}$$

Using Borel's covering theorem (Theorem 13.3, p. 18), $h-h'$ is monotone increasing. Letting h' tend to H we see that $h-H$ is monotone increasing. Similarly for $H-h'$. Again, for $0 < u-x \leqslant \delta_1(x)$, with u in $[a, b]$,

$$\begin{aligned}H(u)-H(x) &= \{h(u)-h(x)\}-\{h(u)-H(u)\}+\{h(x)-H(x)\}\\ &\geqslant f(x)\{g(u)-g(x)\}-\{h(u)-h'(u)\}+\{h(x)-h'(x)\}\end{aligned}$$

$$\begin{aligned}H(u)-H(x) &= \{h'(u)-h'(x)\}+\{H(u)-h'(u)\}-\{H(x)-h'(x)\}\\ &\leqslant f(x)\{g(u)-g(x)\}+\{h(u)-h'(u)\}-\{h(x)-h'(x)\}\end{aligned}$$

$$H(u)-H(x)-f(x)\{g(u)-g(x)\}| \leqslant \{h(u)-h'(u)\}-\{h(x)-h'(x)\}$$

A similar result holds for $0 < x-t \leqslant \delta_1(x)$, with t in $[a, b$ so that

$$V(H-fH_g; [a, b]) \leqslant h(b)-h'(b) < \epsilon$$

As $\epsilon > 0$ is arbitrary we see that H is variationally equivalent to f, g in $[a, b]$.

Conversely, if H is variationally equivalent to f, g in $[a, b]$, with $H(a) = 0$, and using finitely additive χ, then writing $\chi(u, v)$ as $\chi(v)-\chi(u)$, with $\chi(a) = 0$, we see that from the definitions of Section 24 (p. 39), $H+\chi$ is a Ward major and $H-\chi$ a Ward minor function of f, g in $[a, b]$. Further,

$$\{H(b)+\chi(b)\}-\{H(b)-\chi(b)\} = 2\chi(b)$$

and can be as small as we please, so that H is the Ward integral of f, g.

If in (50.2; 50.3) we replace $h(u)-h(x)$, $h(x)-h(t)$ respectively by $h(x, u)$ and $h(t, x)$, where h is finitely superadditive, then the corresponding Ward integral is also equivalent to the Riemann-complete integral. In this case we can take h continuous when g is, by using Theorem 24.2 (p. 41). Further we can suppose that (50.1) is true except possibly in a set of x-variation zero, with Dh never $-\infty$, by use of Theorem 31.3 (p. 62).

It is not possible in a small book like this to assess properly the great contributions of W. H. Young to the theory of integration, as so much has been absorbed into the Lebesgue integration, but the main idea seems to be the use of multiple sequences monotone in each variable separately, covering the Stieltjes form as well as the ordinary form.

Also, no reference is made to the theory of measure, apart from the variation, and Theorem 49.1 (p. 121), as a separate book on the theory is being prepared for this series.

Include

9

LINEAR TOPOLOGICAL SPACES, YOUNG'S INEQUALITY AND INTEGRATION

In this chapter we use the preceding theory to construct examples of linear topological spaces, and then we show how integration processes can be defined in such spaces.

non-empty

51. Linear spaces

A set Z of elements z *is a linear space* in the sense of Banach (1932, p. 26), if there are operations of *addition and scalar multiplication*, where the scalars are real or complex numbers a, and where the operations satisfy the axioms

leading to elements of

L1 $z_1+z_2 = z_2+z_1$;

L2 $z_1+(z_2+z_3) = (z_1+z_2)+z_3$;

L3 $z_1+z_2 = z_1+z_3$ implies $z_2 = z_3$;

L4 $a(z_1+z_2) = az_1+az_2$;

L5 $(a_1+a_2)z = a_1z+a_2z$;

L6 $a_1(a_2z) = (a_1a_2)z$;

L7 $1z = z$.

We put

X

$$-z = (-1)z, \quad z_1-z_2 = z_1+(-1)z_2, \quad z/a = a^{-1}z \ (a \neq 0)$$

THEOREM 51.1. *There is a unique (zero) element θ of Z such that*

$\theta+z=$

$$\textit{if z is in Z then } z+\theta = z \tag{51.1}$$

if for some z in Z, $z+z_2=z$, then $z_2=\theta$. also $0.z=\theta$.

$$\textit{if a is a scalar then } a\theta = \theta \tag{51.2}$$

$$a_1z = a_2z \textit{ implies } a_1 = a_2 \textit{ or } z = \theta \tag{51.3}$$

For (51.1) let z_1 be in Z and put $\theta = 0.z_1$. Then by L2, L5,

$$(z+\theta)+1.z_1 = z+(\theta+1.z_1) = z+(0.z_1+1.z_1)$$
$$= z+(0+1)z_1 = z+1.z_1$$

Hence by L1, L3, we have (51.1). Further, if $z+z_2 = z$ then

$z_2 = 0.z$ then similarly $z+z_2=z$; and if

L3 shows that $z_2 = \theta$. For (51.2), we use L6, obtaining

$$a\theta = a(0.z_1) = (a.0)z_1 = 0.z_1 = \theta$$

For (51.3) we use L5, L6, L7, obtaining from the uniqueness of θ,

$$\begin{aligned}(a_1-a_2)z &= a_1z+(-a_2)z = a_1z+(-1)(a_2z)\\ &=1(a_1z)+(-1)(a_1z) = \{1+(-1)\}(a_1z) = 0.a_1z = \theta\end{aligned}$$

If $a_1 \neq a_2$ then by L6, L7, (51.2),

$$z = 1.z = \{(a_1-a_2)z\}/(a_1-a_2) = \theta/(a_1-a_2) = \theta$$

giving (51.3).

Ex. 51.1. Show that the space of real vectors with a fixed finite number of components is a linear space. Similarly for complex vectors.

Ex. 51.2. Show that the set of real sequences, and the set of complex sequences, form linear spaces.

Ex. 51.3. Show that the set of all point functions variationally integrable with respect to a fixed pair of interval functions in a fixed interval, form a linear space.

Ex. 51.4. Show that pairs of interval functions that are variationally integrable in a fixed interval form a linear space.

52. Linear topological spaces

If Z is a linear space, and if V_1, V_2 are sets of elements of Z and a is a scalar, then we denote by

$$z+V_1,\quad V_1+V_2,\quad aV_1$$

the respective sets of points $z+z_1$, z_1+z_2, az_1, for all z_1 in V_1 and all z_2 in V_2. Then *Z is a linear topological space* in the sense of von Neumann (1935, p. 4), if it contains a family $\mathcal{N}$ of subsets such that

T1 θ is in each V of $\mathcal{N}$;

T2 if z is in every V of $\mathcal{N}$ then $z = \theta$;

T3 if V_3, V_4 are in $\mathcal{N}$, there is a V_5 of $\mathcal{N}$ such that $V_5 \subseteq V_3 \cap V_4$;

T4 if V is in $\mathcal{N}$ there is a V_6 of $\mathcal{N}$ such that $V_6+V_6 \subseteq V$;

T5 if V is in $\mathcal{N}$ there is a V_7 of $\mathcal{N}$ such that $aV_7 \subseteq V$ for all $|a| \leqslant 1$;

T6 if z is in Z, and V in $\mathcal{N}$, there is a constant a such that z is in aV.

T7 The space Z is complete.

The last axiom requires the following definitions. A sequence $\{z_j\}$ of elements of Z *is fundamental* if, given V in $\mathcal{N}$, there is an integer J such that for all $j \geqslant J$, $k \geqslant J$, z_j-z_k lies in V. The sequence is *convergent to an element z of Z* if, given V in $\mathcal{N}$, there is an integer J such that for all $j \geqslant J$, z_j-z lies in V. The space *Z is complete* if every fundamental sequence is convergent to an element z of Z.

THEOREM 52.1.

$$2V \subseteq V+V \tag{52.1}$$

Given V in $\mathcal{N}$, there is a sequence $\{V_{8j}\}$ of sets of $\mathcal{N}$, independent of the scalar a, and satisfying

$$\begin{aligned}V_{80} = V,\ V_{8j}+V_{8j} \subseteq V_{8,j-1},\ V_{8j}-V_{8j} \subseteq V_{8,j-1},\\ aV_{8j} \subseteq V_{8,j-1}\ (|a| \leqslant 2, j = 1, 2, \ldots)\end{aligned} \tag{52.2}$$

$$\sum_{n=j}^{k} V_{8n}+V_{8k} \subseteq V_{8,j-1}\ (k \geqslant j) \tag{52.3}$$

For (52.1) we use L5, L7; for

$$2z = (1+1)z = 1z+1z = z+z$$

Replacing the V of T5 by the V_6 of T4, then for all $|a| \leqslant 1$ there is a V_9 of $\mathcal{N}$ such that

$$\begin{aligned}aV_9 \subseteq V_6,\ 2aV_9 \subseteq 2V_6 \subseteq V_6+V_6 \subseteq V,\\ -V_9 =(-1)V_9 \subseteq V_6,\ \ V_9-V_9 \subseteq V_6+V_6 \subseteq V\end{aligned}$$

For $V = V_{8,j-1}$ then the V_9 is V_{8j}

From (52.2) we obtain (52.3)

A point *z is a limit-point of the set W of points of Z if*, given V in $\mathcal{N}$, there is a $z' \neq z$ such that z' is in both W and $z+V$. The closure $\overline{W}$ of W is the set of all points and limit-points of W.

THEOREM 52.2. *If V is in $\mathcal{N}$, then*

$$\overline{V}_{81} \subseteq V,\quad -\overline{V}_{81} \subseteq V$$

For let z be a limit-point of V_{81}. There is then a $z' \neq z$ and contained in V_{81} and in $z+V_{81}$. By (52.2),

$$z'-z \in V_{81}, \quad z = z'-(z'-z) \in V_{81}-V_{81} \subseteq V,$$
$$-z = (z'-z)-z' \in V_{81}-V_{81} \subseteq V$$

Ex. 52.1. A linear space Z with zero element θ is said to be *normed* if to each element z of Z there corresponds a real number $||z||$, called the *norm of* z, such that

N1 $||z|| = 0$ if and only if $z = \theta$;

N2 $||az|| = |a| \,.\, ||z||$ for each scalar a;

N3 $||z_1+z_2|| \leqslant ||z_1||+||z_2||$

First prove that $||z|| \geqslant 0$ by considering all three axioms and using $-z$; and then show that z is a linear topological space, if $\mathscr{N}$ is the family of sets of Z that are defined by $||z|| < \epsilon$ for some $\epsilon > 0$, and if the resulting space is complete.

53. Young's inequality

We now use the preceding theory to obtain examples of linear topological spaces. First, we need a Stieltjes transformation theorem.

Theorem 53.1. *Let φ, ψ be monotone increasing inverse functions as in Theorem* 23.1 (p. 35). *If φ is continuous and the first integral exists, then*

$$(V)\int_{\psi(a)}^{\psi(b)} f\,d\varphi = (V)\int_a^b f(\psi(y))\,dy \quad (b > a \geqslant 0) \tag{53.1}$$

In particular,

$$(V)\int_{\psi(a)}^{\psi(b)} x\,d\varphi = (V)\int_a^b \psi(y)\,dy \tag{53.2}$$

In Theorem 23.1 (p. 35) we take

$$h_l(u, v) = f(v)\{\varphi(v)-\varphi(u)\}, \quad h_r(u, v) = f(u)\{\varphi(v)-\varphi(u)\}$$

Then by Theorem 23.2 (p. 37)

$$h_l{}^*(y, z) = f(\psi(z))(z-y), \quad h_r{}^*(y, z) = f(\psi(y+))(z-y)$$

and **h**, **h*** satisfy the conditions of Theorem 23.1.
The points x where

$$\psi(x) < \psi(x+)$$

$(\eta_j = \varphi(\xi_j), \psi(\eta_j) = \xi_j^*)$

Prove $\int_0^a \varphi(x)dx + \int_0^{\varphi(a)} \psi(y)dy = a\varphi(a)$. $\sum \varphi(\xi_j)(x_j - x_{j-1}) + \sum \psi(\eta_j)(y_j - y_{j-1})$

ok if φ const.

So $\int_a^b \varphi\,dx + \int_{\varphi(a)}^{\varphi(b)} \psi\,dy = [x\varphi(x)]_a^b =$ $\varphi(\xi_j^*)$(length of interval where φ const.) $\times$ jump of ψ.

form an at most countable set, and so a set of x-variation zero, so that Theorem 31.3, p. 62, (31.17) implying (31.16), completes the proof.

Theorem 53.2. (W. H. Young). *Let φ, ψ be monotone increasing inverse functions as in Theorem* 23.1 (p. 35). *If φ is continuous and $a > 0$, $b > 0$, we have*

$$(V)\int_0^a \varphi(x)\,dx + (V)\int_0^b \psi(x)\,dx \geqslant ab \tag{53.3}$$

We write the respective integrals as $\Phi(a)$, $\Psi(b)$. *If equality in* (53.3) *then*

$$\varphi(a) = b \tag{53.4}$$

In particular we have

$$\frac{a^p}{p} + \frac{b^q}{q} \geqslant ab \left(a \geqslant 0, b \geqslant 0, p > 1, \frac{1}{p} + \frac{1}{q} = 1\right) \tag{53.5}$$

with equality when $a^p = b^q$;

$$a \log a + e^{b-1} \geqslant ab \quad (a > 0, b \geqslant 0) \tag{53.6}$$

with equality when $a \geqslant e^{-1}$ *and* $b = \log a + 1$

From Theorem 53.1 (53.1) we can write the left-hand side of (53.3) as

$$(V)\int_0^a \varphi(x)\,dx + (V)\int_0^{\psi(b)} x\,d\varphi(x) \tag{53.7}$$

Suppose first that $\psi(b) \geqslant a$. Then (53.7) is

$$\geqslant (V)\int_0^a \varphi(x)\,dx + (V)\int_0^a x\,d\varphi(x) + (V)\int_a^{\psi(b)} a\,d\varphi(x),$$

Now x and φ satisfy Theorem 33.1 (33.2), p. 69, so that by the integration by parts and Theorem 23.2 [redacted], p. 37, (53.7) is

$$\geqslant a\varphi(a) - 0\varphi(0) + a\{b-\varphi(a)\} = ab$$

If $a > \psi(b)$ then (53.7) is

$$\geqslant (V)\int_0^{\psi(b)} \varphi(x)\,dx + (V)\int_0^{\psi(b)} x\,d\varphi(x) + (V)\int_{\psi(b)}^a \varphi(\psi(b))\,dx$$

[illegible] $\varphi(x)dx + \int_0^{\varphi(a)} \psi(y)dy + \{b - \varphi(a)\}\psi(\varphi(a))$ [Take a to left of any interval of constancy of φ, & then $\psi(\varphi(a)) = a$. For if a larger then $\int_0^a \varphi dx$ larger.

Kestelman: graph of e^{t-1} lies above its tangents: in particular, at $t_0 = 1 + \log a$, one gets

Again by Theorems 33.1; 23.2, (53.7) is

$$\geqslant \psi(b)\varphi(\psi(b)) - 0\varphi(0) + b\{a - \psi(b)\} = ab$$

Thus (53.3) is true. If equality occurs in (53.3) then for $\psi(b) > a$ we have

$$(V)\int_a^{\psi(b)} (x-a)\,\mathrm{d}\varphi(x) = 0$$

and φ is constant in a to $\psi(b)$. Hence by Theorem 23.2, (53.4) follows. If equality in (53.3) and $\psi(b) < a$, then

$$(V)\int_{\psi(b)}^{a} \{\varphi(x) - b\}\,\mathrm{d}x = 0$$

and $\varphi(x) = b$ in a to $\psi(b)$, giving (53.4) again. Finally, if $\psi(b) = a$ then Theorem 23.2 gives (53.4).

To prove (53.5) we can assume $a > 0, b > 0, \varphi(x) = x^{p-1}$. As $p-1 > 0$ we have $\varphi(0) = 0$, and φ is continuous and monotone increasing, and gives

$$\psi(y) = y^{1/(p-1)},\ \Phi(x) = x^p/p,\ \Psi(y) = y^q/q,$$

where

$$q = 1 + 1/(p-1) = p/(p-1),\quad 1/p + 1/q = 1$$

For (53.6) we take

$$\varphi(x) = 0(0 \leqslant x \leqslant e^{-1}),\ \log x + 1(x > e^{-1})$$
$$\psi(0) = 0,\ \psi(0+) = e^{-1},\ \psi(y) = e^{y-1}(y > 0)$$
$$\Phi(x) = 0(0 \leqslant x \leqslant e^{-1}),\ x\log x + e^{-1}(x > e^{-1})$$
$$\Psi(y) = e^{y-1} - e^{-1}$$

Thus for $a \geqslant e^{-1}, b > 0$, (53.3) gives (53.6), and equality and (53.4) gives $b = \log a + 1$. If $0 < a < e^{-1}, b > 0$, (53.3) gives

$$ab \leqslant 0 + e^{b-1} - e^{-1}$$

As $a \log a$ has a minimum at $a = e^{-1}$ we have (53.6) again, with strict inequality; and also the case $b = 0$ follows.

THEOREM 53.3. *Let* $\mathrm{f} \geqslant 0, g \geqslant 0$ *be point functions,* $\alpha > 0$, $\beta > 0$ *be constants,* $\mathrm{h}_s \geqslant 0(s = 1, r)$ *be interval functions, and* $\varphi, \psi, \Phi, \Psi$ *be as in Theorem 53.2. If* fg, $\Phi(\alpha\mathrm{f})$, $\Psi(\beta\mathrm{g})$ *are*

---

all variationally integrable with respect to **h** *in an elementary set* E *then*

$$(V)\int_E \mathrm{fg}\,d\mathbf{h} \leqslant (V)\int_E \Phi(\alpha\mathrm{f})\,d\mathbf{h} + (V)\int_E \Psi(\beta\mathrm{g})\,d\mathbf{h} \qquad (53.8)$$

If there are values α, β *such that the right-hand side of* (53.8) *is* 1 *then,*

$$(V)\int_E \mathrm{fg}\,d\mathbf{h} \leqslant \alpha^{-1}\beta^{-1} \qquad (53.9)$$

In particular, if $\mathrm{p} > 1, \mathrm{p}^{-1} + \mathrm{q}^{-1} = 1$, *we have Hölder's inequality*

$$(V)\int_E \mathrm{fg}\,d\mathbf{h} \leqslant \left[(V)\int_E \mathrm{f}^{\mathrm{p}}\,d\mathbf{h}\right]^{1/\mathrm{p}} \left[(V)\int_E \mathrm{g}^{\mathrm{q}}\,d\mathbf{h}\right]^{1/\mathrm{q}} \qquad (53.10)$$

If equality occurs in (53.8; 53.10) *then, apart from a set of* x *of* **h***-variation zero we have, respectively,* $\varphi(\alpha f) = \beta\mathrm{g}$, *and*

$$\mathrm{f}^{\mathrm{p}}(\mathrm{x})/(V)\int_E \mathrm{f}^{\mathrm{p}}\,d\mathbf{h} = \mathrm{g}^{\mathrm{q}}(\mathrm{x})/(V)\int_E \mathrm{g}^{\mathrm{q}}\,d\mathbf{h} \qquad (53.11)$$

To obtain (53.8) we integrate (53.3) with respect to **h** in *E*, with $a = \alpha f$, and $b = \beta g$. Then (53.9) is easy, and (53.5; 53.9) with

$$\alpha = \left[(V)\int_E f^p\,\mathrm{d}\mathbf{h}\right]^{-1/p},\quad \beta = \left[(V)\int_E g^q\,\mathrm{d}\mathbf{h}\right]^{-1/q}$$

give (53.10). The cases of equality follow from Ex. 31.1 (p. 64), and Theorem 53.2 (53.4).

THEOREM 53.4. *In Theorem* 53.3 (53.10) *we need not assume that* fg *is variationally integrable with respect to* **h** *in* E, *if* f *is continuous, non-negative, and variationally integrable in* E.

Drop continuity, for see P Y Lee, JLMS 40 (1965) 33[illegible]-3[illegible]

Let us put

$$f_n = \min(f, n),\quad g_n = \min(g, n)$$

Then by Theorem 25.4 (p. 46), $f_n{}^p$ and $g_n{}^q$ are variationally integrable with respect to **h** in E, and so are **h**-measurable, by Theorem 39.2 (p. 94). Then $f_n g_n$ is **h**-measurable, and is bounded by n^2, so that it is variationally integrable with respect to **h** in E. Letting $n \to \infty$ and using Theorems 36.1 (p. 82) and 53.3 (53.10) for f_n, g_n, we obtain what we require. continuity can be dropped, see Th. 53.4.

(p) →

THEOREM 53.5. *Let* $p > 1$, *and let* $f \geqslant 0, g \geqslant 0$ *be point functions, and* $h_s \geqslant 0(s = 1, r)$ *be interval functions, such that* f^p *and* g^{q} *are variationally integrable with respect to* **h** *in* E. *If either* $f(f+g)^{p-1}$ *and* $g(f+g)^{p-1}$ *are variationally integrable in* E *with respect to* **h**, *or* **h** *is continuous and variationally integrable in* E, *then we have Minkowski's inequality*

$$\left[(V)\int_E (f+g)^p\, d\mathbf{h}\right]^{1/p} \leqslant \left[(V)\int_E f^p\, d\mathbf{h}\right]^{1/p} + \left[(V)\int_E g^p\, d\mathbf{h}\right]^{1/p}$$

From Theorem 53.2 (53.5) we have

$$\alpha\beta(a_1b_1+a_2b_2) \leqslant p^{-1}\alpha^p(a_1{}^p+a_2{}^p)+q^{-1}\beta^q(b_1{}^q+b_2{}^q)$$

$$a_1b_1+a_2b_2 \leqslant (a_1{}^p+a_2{}^p)^{1/p}(b_1{}^q+b_2{}^q)^{1/q}$$

$$(a_j \geqslant 0, b_j \geqslant 0, j = 1, 2)$$

$$(a+b)^p = a(a+b)^{p-1} + b(a+b)^{p-1} \leqslant (a^p+b^p)^{1/p}2^{1/q}(a+b)^{p/q}$$

$$a+b \leqslant 2^{1/q}(a^p+b^p)^{1/p}$$

$$a_1b_1+a_2b_2 \quad (a_1{}^p+a_2{}^p)^{1/p}(b_1{}^q+b_2{}^q)^{1/q}$$

$$(a_j \geqslant 0, b_j \geqslant 0, j = 1,2)$$

$$(a+b)^p = a(a+b)^{p-1}+b(a+b)^{p-1} \leqslant (a^p+b^p)^{1/p}2^{1/q}(a+b)^{p/q}$$

$$a+b \leqslant 2^{1/q}(a^p+b^p)^{1/p}$$

Thus an argument similar to that in Theorem 53.4 will show that when **h** is continuous, non-negative, and variationally integrable in E, the variational integrability of $f(f+g)^{p-1}$ and $g(f+g)^{p-1}$ with respect to **h** in E, follows from the integrability of f^p, g^p. Further, under either group of alternative hypotheses,

$a(a+b)^{p-1} = a(a+b)^{p/q} \leq a^p/p + (a+b)^p/q \leq a^p/p + 2^{p/q}(a^p+b^p)/q$ $(a \geq 0, b \geq 0)$. | $b(a+b)^{p-1} \leq b^p/p + 2^{p/q}(a^p+b^p)/q$ $(a \geq 0, b \geq 0)$.

by Hölder's inequality,

$$(V)\int_E (f+g)^p\,\mathbf{dh} = (V)\int_E f(f+g)^{p-1}\,\mathbf{dh} + (V)\int_E g(f+g)^{p-1}\,\mathbf{dh} \leqslant \left[(V)\int_E f^p\,\mathbf{dh}\right]^{1/p} \times \left[(V)\int_E (f+g)^p\,\mathbf{dh}\right]^{1/q} + \left[(V)\int_E g^p\,\mathbf{dh}\right]^{1/p}\left[(V)\int_E (f+g)^p\,\mathbf{dh}\right]^{1/q}$$

Hence dividing by the extra factor we obtain Minkowski's inequality. Note that the condition for equality is obtained from (53.11) and gives

$$f^p(x)/(V)\int_E f^p\,\mathbf{dh} = g^p(x)/(V)\int_E g^p\,\mathbf{dh}$$

except possibly in a set of x of **h**-variation zero, so that f is a multiple of g almost everywhere.

54. Mean convergence

A more sophisticated kind of convergence of sequences of functions is dealt with in the following theorem.

THEOREM 54.1. *Let* $h_s \geqslant 0(s = 1, r)$ *be interval functions, and let* $p > 0$. *If* $\{f(j; x)\}$ *is a sequence of point functions such that, given* $\epsilon > 0$, *there is an integer* $k(\epsilon)$ *with the property that*

$$(V)\int_E |f(j;x)-f(k;x)|^p\, d\mathbf{h} \tag{54.1}$$

exists for $j \geqslant k(\epsilon)$, $k \geqslant k(\epsilon)$, *and is less than* ϵ, *then there are a point function* f *and a subsequence* $\{j_n\}$ *of the integers such that*

$$\lim_{n\to\infty} f(j_n; .) = f \tag{54.2}$$

except in a set X *with*

$$V(\mathbf{h}; E; X) = 0 \tag{54.3}$$

and for each integer j $\geqslant$ k(ϵ)

$$(\mathrm{V})\int_{\mathrm{E}} |\mathrm{f}(\mathrm{j};\,.)-\mathrm{f}|^{\mathrm{p}}\,d\mathbf{h} \tag{54.4}$$

exists and is not greater than ϵ. *Further, if* g *also satisfies the condition for* f *in* (54.4) *then* f = g *except in a set satisfying* (54.3).

If $\{f(j;\,.)\}$ is a sequence of point functions satisfying the hypotheses of the theorem we say that the sequence *converges in mean with index p, relative to* **h** *and E*. If $\{f(j;\,.)\}$ satisfies the conclusion of the theorem we say that the sequence *converges in mean to f, with index p, relative to* **h** *and E*.

To each integer n corresponds an integer j_n such that

$$(V)\int_{\mathrm{E}} |f(j;\,.)-f(k;\,.)|^p\,d\mathbf{h} < 2^{-2np}\ (j \geqslant j_n,\ k \geqslant j_n)$$

Clearly we can suppose that $j_n < j_{n+1}$ for $n = 1, 2, \ldots$. Then by Theorem 30.1 (30.3), p. 57, if X_{1n} is the set where

$$|f(j_n;\,.)-f(j_{n+1};\,.)| \geqslant 2^{-n}$$

then

$$V(\mathbf{h};\,\mathrm{E};\,X_{1n}) < 2^{-np}$$

It follows from Theorem 28.1 (p. 49) that if X_{2N} is the union of $X_{1N}, X_{1,N+1}, \ldots,$

$$V(\mathbf{h};\,\mathrm{E};\,X_{2N}) < 2^{-Np}/(1-2^{-p}) \tag{54.5}$$

If x is not in the set X_{2N} then

$$\sum_{n=N}^{\infty} |f(j_n;\,.)-f(j_{n+1};\,.)| < \sum_{n=N}^{\infty} 2^{-n} = 2^{-N+1}$$

so that

$$f = \lim_{n\to\infty} f(j_n;\,.)$$

exists. Thus

$$X \subseteq \bigcap_{N=1}^{\infty} X_{2N},\ V(\mathbf{h};\,E;\,X) \leqslant V(\mathbf{h};\,\mathrm{E};\,X_{2N})$$

and (54.5) gives (54.3). The integrand is non-negative, so that by Fatou's lemma (Theorem 37.1 (37.3), p. 85) we have

$$(V)\int_E |f(j;\,.)-f|^p\,d\mathbf{h} = (V)\int_E |f(j;\,.)-\lim_{n\to\infty} f(j_n;\,.)|^p\,d\mathbf{h}$$

$$\leqslant \liminf_{n\to\infty}(V)\int_E |f(j;\,.)-f(j_n;\,.)|^p\,d\mathbf{h} \leqslant \epsilon\ (j \geqslant k(\epsilon))$$

This gives (54.4). When for f we can substitute g, then by Theorem 37.1 (37.3), p. 85,

$$(V)\int_E |f-g|^p\,d\mathbf{h} \leqslant \liminf_{n\to\infty}(V)\int_E |f(j_n;\,.)-g|^p\,d\mathbf{h} \leqslant \epsilon$$

for each $\epsilon > 0$, so that

$$(V)\int_E |f-g|^p\,d\mathbf{h} = 0$$

By Ex. 31.1 (p. 64), $f = g$ except in a set satisfying (54.3).

Theorem 54.2. *Let* $\mathrm{h_s} \geqslant 0(\mathrm{s} = 1, \mathrm{r})$ *be interval functions, and let* $\mathrm{p} > 1, \mathrm{p}^{-1}+\mathrm{q}^{-1} = 1$. *If* g $\geqslant$ 0 *is a point function, and* $\{\mathrm{f}(\mathrm{j};\,\mathrm{x})\}$ *a sequence of point functions, such that* f(j; .)g, g^{q}, *and* $\mathrm{f}(\mathrm{j};\,.)-\mathrm{f}(\mathrm{k};\,.)|^{\mathrm{p}}$ (j, k = 1, 2, ...) *are all variationally integrable with respect to* **h** *in an elementary set* E, *with* (54.1) *less than a given* $\epsilon > 0$, *for* j $\geqslant$ k(ϵ), k $\geqslant$ k(ϵ), *and if* f *is the function of Theorem* 54.1, then

$$\mathit{lim}_{\mathrm{n}\to\infty}\,(\mathrm{V})\int_{\mathrm{E}} \mathrm{f}(\mathrm{j};\,.)\mathrm{g}\,d\mathbf{h} = (\mathrm{V})\int_{\mathrm{E}} \mathrm{fg}\,d\mathbf{h}$$

By Theorems 19.1 (p. 27); 25.2 (p. 45); and 53.3 (53.10) we have the existence of the first integral in (54.6) and the result

$$(V)\int_E |f(j;\,.)-f(k;\,.)|g\,d\mathbf{h} \leqslant$$

$$\left[(V)\int_E |f(j;\,.)-f(k;\,.)|^p\,d\mathbf{h}\right]^{1/p}\left[(V)\int_E g^q\,d\mathbf{h}\right]^{1/q} \tag{54.6}$$

Let us put

$$t(j,k;.) = \max\{f(j;.)-f(k;.);0\}$$
$$=\tfrac{1}{2}\{|f(j;.)-f(k;.)|+f(j;.)-f(k;.)\}$$

t Then $\cancel{f}g$ is variationally integrable with respect to **h** in E, and from (54.6),

$$(V)\int_E tg\,\mathbf{dh} \leqslant$$
$$\left[(V)\int_E |f(j;.)-f(k;.)|^p\,\mathbf{dh}\right]^{1/p}\left[(V)\int_E g^q\,\mathbf{dh}\right]^{1/q} \qquad (54.7)$$

Also $t(j;j_n;.)$ tends as $n\to\infty$, to

$$\max\{f(j;.)-f;0\}$$

which by (54.7) and Theorem 37.1 (37.3), p. 85, is, when multiplied by g, variationally integrable with respect to **h** in E. A similar result holds for

$$\min\{f(j;.)-f;0\}$$

so that $\{f(j;.)-f\}g$, and so fg, are variationally integrable with respect to **h** in E. Hence as in (54.6),

$$\left|(V)\int_E f(j;.)g\,\mathbf{dh}-(V)\int_E fg\,\mathbf{dh}\right| = \left|(V)\int_E \{f(j;.)-f\}g\,\mathbf{dh}\right|$$
$$\leqslant (V)\int_E |f(j;.)-f|g\,\mathbf{dh}$$
$$\leqslant \left[(V)\int_E |f(j;.)-f|^p\,\mathbf{dh}\right]^{1/p}\left[(V)\int_E g^q\,\mathbf{dh}\right]^{1/q}$$

and Theorem 54.1 completes the result.

Ex. 54.1. In Theorem 54.1 we can replace $|\ldots|^p$ by $\Phi(\ldots)$ if for some $\epsilon>0$, $\delta>0$, $p>0$, we have

$$\Phi(x) \geqslant \epsilon\{\min(x,\delta)\}^p \quad (x>0)$$

Ex. 54.2. Under the hypotheses of Theorem 54.1, if X_{3jk} is the set of x where

$$|f(j;x)-f(k;x)| \geqslant \epsilon$$

prove that $V(\mathbf{h};E;X_{3jk})$ tends to 0 as j,k tend to infinity.

Such a sequence $\{f(j;x)\}$ is usually said to be *fundamental in measure*, when **h** is the length of the interval in question.

$$(V)\int_E |f(j;.)-f(k;.)|^p\,\mathbf{dh}$$
$$=V\Big((V)\int ;E\Big) \geqslant V\Big((V)\int ;E;X_{3jk}\Big) \geqslant \epsilon V(\mathbf{h};E;X_{3jk})\Big)$$

55. The complete normed linear space L^p $(p>1)$

Let us fix the interval functions $h_s \geqslant 0 (s = l, r)$ and the elementary set E, in this section. We then denote by L^p the space of all functions f such that f and $|f|^p$ are variationally integrable with respect to **h** in E. By Theorem 53.5 p.134 (Minkowski's inequality), L^p is linear. We define

$$||f|| = \left[(V)\int_E |f|^p\,\mathbf{dh}\right]^{1/p}$$

Then N2 and N3 of Ex. 52.1 p.130 are satisfied. If $||f||=0$ then by Ex. 31.1 (p. 64) we have $f=0$ except in a set satisfying (54.3), so that for a zero in the space L^p we have to use the functions f that are zero except in a set satisfying (54.3), called *null functions*. Thus we are led to change the definition of L^p, taking the elements of L^p to be families (f), where (f) is the family of all functions g such that $g=f$ except in a set satisfying (54.3), i.e. $g-f$ is a null function. By Theorem [illegible] L^p is still linear, and by Theorem 31.3 p.62, (31.17) implying [illegible], we have $||g||=||f||$ if $g-f$ is a null function. Thus we can define $||(f)||=||g||$, for arbitrary g in (f), the norm satisfying N1, N2, N3 of Ex. 52.1. Finally, by Theorem 54.1 p.135 the space L^p is complete, so that L^p can be written as a linear topological space.

More generally, we can use the Φ, Ψ of Theorem 53.2 p.131, in place of $|f|^p$. Then in Ex. 52.1, N1 can be made true by a device similar to the above, when

$$||f||^* = (V)\int_E \Phi(|f|)\,\mathbf{dh} \qquad (55.1)$$

Completeness is ensured by Ex. 54.1 p.138 when Φ satisfies the condition given there. But N2, N3, involving the linearity of the space, require more care. The definition (55.1) has to be replaced by

$$||f|| = \sup (V) \int_E |fg| \,d\mathbf{h} \qquad (55.2)$$

for all g such that

$$(V) \int_E \Psi(|g|) \,d\mathbf{h} \leqslant 1$$

Then L_Φ is the (Orlicz) space of all f such that (55.2) is finite. It can be proved that when $||f||^*$ is finite so is $||f||$; that when $||f||$ is finite so is $||\alpha f||^*$, for some constant $\alpha > 0$; that L_Φ satisfies N1, N2, N3; and that L_Φ is complete. See, for example, Zaanen (1953, pp. 78–83).

*56. The Riemann-complete integral for functions with values in a linear topological space

We now return to the general case, where we are concerned with the following kind of convergence. To each $\mathbf{A}$ complete in E let there correspond a set $S(\mathbf{A})$ of points of the linear topological space Z, satisfying

$$S(\mathbf{A}) \subseteq S(\mathbf{A}_1) \text{ if } \mathbf{A} = \mathbf{A}.\mathbf{A}_1 \qquad (56.1)$$

Then $S(.)$ *is a fundamental set relative to sets complete in* E if, given V in $\mathscr{N}$, there is an $\mathbf{A}_2$ complete in E such that

$$S(\mathbf{A}_2) - S(\mathbf{A}_2) \subseteq V \qquad (56.2)$$

A point z *of* Z *is the limit of* $S(.)$, *relative to sets complete in* E, if, given a set V of $\mathscr{N}$, there is a set $\mathbf{A}_3$ complete in E such that

$$S(\mathbf{A}_3) \subseteq z + V \qquad (56.3)$$

The use of T2, T4 ,T5 (or T2 and Th. 52.1(52.2), p.129) ensures that $S(.)$ cannot have more than one limit. To show that a set $S(.)$ satisfying (56.3) is fundamental we need only observe that by Theorem 52.1 (52.2) p.129 if $z_j - z$ lies in V_{81} for $j = 1, 2$, then

$$z_1 - z_2 = (z_1 - z) - (z_2 - z) \subseteq V_{81} - V_{81} \subseteq V$$

Thus replacing V by V_{81} in (56.3) we obtain (56.2).

The space Z can be called *complete relative to* $\mathbf{A}$, if every fundamental set relative to $\mathbf{A}$ has a limit relative to $\mathbf{A}$. We need to strengthen T7 to become the following.

T8 Z is complete relative to $\mathbf{A}$.

THEOREM 56.1. T8 *implies* T7. *If there is a fixed sequence* $\{V_{1j}\}$ *of sets of* $\mathscr{N}$ *with the property that for each* V *in* $\mathscr{N}$ *there is an integer* j *such that* V_{1j} *is in* V, *then* T7 *implies* T8.

If $S(\mathbf{A})$ is the set of values z_j for some k and all $j \geqslant k$, where k varies with $\mathbf{A}$, and where $\{z_j\}$ is fundamental, then $S(.)$ is a fundamental set, so that by T8 it has a limit, say z. Then $\{z_j\}$ converges to z and T7 is true. Conversely, if T7 is true, and if there is a fixed sequence $\{V_{1j}\}$ with the given property, we have to show that T8 is true. By T4 there is a V_{2j} such that $V_{2j} + V_{2j} \subseteq V_{1j}$, and then by hypothesis there is a $k > j$ such that

$$V_{1k} \subseteq V_{2j}, \quad V_{1k} + V_{1k} \subseteq V_{1j}$$

By taking a subsequence if necessary, we can therefore assume that

$$V_{1j} + V_{1j} \subseteq V_{1,j-1} (j = 1, 2, \ldots) \qquad (56.4)$$

Let $\mathbf{A}_{4j}$ be the $\mathbf{A}_{2j}$ for (56.2) when V is V_{1j}, and let z_j be a point of $S(\mathbf{A}_{4j})$, so that by (56.1), z_k is also in $S(\mathbf{A}_{4j})$ for each $k > j$. Then

$$z_j - z_k \in S(\mathbf{A}_{4j}) - S(\mathbf{A}_{4j}) \subseteq V_{1j} \; (k > j)$$

so that $\{z_j\}$ is fundamental, and so convergent by T7 to some z of Z. Hence there is a $k \geqslant j$ such that by (56.1),

$$\begin{aligned} z_k - z \in V_{1j}, \; S(\mathbf{A}_{4j}) - z &= S(\mathbf{A}_{4j}) - z_k + z_k - z \\ &\subseteq S(\mathbf{A}_{4j}) - S(\mathbf{A}_{4j}) + V_{1j} \subseteq V_{1j} + V_{1j} \subseteq V_{1,j-1} \end{aligned}$$

Hence $S(.)$ has the limit z, relative to complete sets in E, so that T7 gives T8.

Let $\mathbf{h}$ be a pair of functions of intervals $[u, v)$, with values in Z, and let $\mathbf{A}$ be complete in E. We consider sets $S(\mathbf{h}; \mathbf{A}; E)$ of sums

$$(\mathscr{D}) \sum h_s(u, v)$$

over divisions $\mathscr{D}\{u, v\}$ of E formed from $\mathbf{A}$. If there exists a limit z of $S(\mathbf{h}; . ; E)$, relative to sets complete in E, we call it the Riemann-complete integral $RC(\mathbf{h}; E)$ of $\mathbf{h}$ in E. Straightforward translation of the previous theory from functions with real or complex values to functions with values in z, can be achieved for Theorems 19.1 (p. 27); 21.1 p.31, 21.2 (p. 32); 23.1 p.35, 23.3 (p. 37); 46.1 (p. 113); 48.1 (p. 118), by replacing circular neighbourhoods by $z + V$, for V in $\mathscr{N}$, and using Theorem 52.1 (52.2; 52.3) p.129 where necessary.

Phillips (1940), Price (1940) and Rickart (1942), have considered Lebesgue-type integration for functions with values in general spaces, using divisions containing an infinity of elements, and an integral of Getchell type, and as usual in theory dealing with Lebesgue-type integrals they presuppose that a measure is already given for general sets. Their divisions contain an infinity of elements to ensure that their final integral is completely additive, on the mistaken idea that complete additivity of the measure is necessary in order that the principal theorems of Lebesgue theory should hold. But the foregoing pages show that complete additivity is an unnecessary hypothesis, and that finite additivity is enough. Further, divisions with an infinity of elements can lead to unsatisfactory results as is shown in Ex. 56.1.

We now consider an operator T defined for all z in Z, with values in a linear topological space Y. T is *distributive* if for each z, z^* in Z, and each scalar a, b,

$$T(az+bz^*) = aTz+bTz^*$$

The operator is *continuous* if, given z in Z and V_Y in $\mathscr{N}_Y$, the $\mathscr{N}$ for Y, there is a V_3 in $\mathscr{N}$ such that if z' is in $z+V_3$ then Tz' is in $Tz+V_Y$. The operator is *closed* if $y = Tz$ whenever a set $S(.)$ has a limit z in Z and the set $TS(.)$ of values Tz, for z in $S(.)$, has a limit y in Y. Thus a continuous operator is closed, but a closed operator need not be continuous.

THEOREM 56.2. If *the operator* T *is distributive and continuous, and if* **h** *is Riemann-complete integrable in* E, *then* $\mathbf{Th} = \{\mathrm{Th_l}, \mathrm{Th_r}\}$ is *Riemann-complete integrable in* E *satisfying*

$$\mathrm{TRC}(\mathbf{h}; \mathrm{E}) = \mathrm{RC}(\mathrm{T}\mathbf{h}; \mathrm{E}) \tag{56.5}$$

If T *is distributive and closed, and if both sides of* (56.5) *exist, then they are equal.*

In both cases T is distributive, so that if $\mathscr{D}_1$ is a division of E,

$$Tz(\mathscr{D}_1) = y(\mathscr{D}_1)$$

where

$$z(\mathscr{D}_1) = (\mathscr{D}_1)\sum h_s, \quad y(\mathscr{D}_1) = (\mathscr{D}_1)\sum Th_s \tag{56.6}$$

In the first case T is also continuous. Thus if $z = \mathrm{RC}(\mathbf{h}; \mathrm{E})$, we can take $z(\mathscr{D}_1)$ for z', where $\mathscr{D}_1$ comes from a suitable $\mathbf{A}_5$ complete in E, so that z' lies in $z+V_3$. From (56.6) the corresponding $y(\mathscr{D}_1)$ lies in $Tz+V_Y$, and $T\mathbf{h}$ is Riemann-complete integrable to Tz, giving (56.5).

In the second case, $S(\mathbf{h}; .; E)$ has the limit $z = RC(\mathbf{h}; E)$, and $S(T\mathbf{h}; .; E)$ has the limit $y = RC(T\mathbf{h}; E)$, both relative to complete sets in E. By (56.6) the second S-set is the application of T to the first S-set, so that as T is closed we again have (56.5).

Ex. 56.1. Let

$$h(u, v) = \begin{cases} 0 (0 \leqslant u < v < 1) \\ 1 (0 \leqslant u < v = 1) \end{cases}$$

which is finitely additive and independent of l, r. Then show that

$$RC(h; 0, 1) = 1 = h(0, 1)$$

but if, for example, we use the division $\mathscr{D}_2$ of elements $[1-2^{-j}, 1-2^{-j-1})$ for $j = 0, 1, 2, \ldots$, the sum is zero, and similarly for all divisions of $[0, 1]$ that do not include an interval $[1-\epsilon, 1)$ for some $\epsilon > 0$. Thus the Rickart integral of h is 0.

More generally, it can be proved that the Rickart integral of a finitely additive function gives the completely additive part of the function, if it exists, so that we have lost what may be an important part of the function.

*57. Functions of bounded variation

Let **A** be complete in E. Then by $P(\mathbf{h}; \mathbf{A}; E)$, *the variation set of* **h** *in* E, *using* **A**, we denote the set of all sums

$$(\mathscr{Q})\sum h_s(u, v)$$

obtained from finite collections $\mathscr{Q}$ of disjoint intervals $[u, v)$ in E from **A**, where as usual, $s = l$ if $[u, v)$ is from $\mathscr{L}$, and otherwise $s = r$. By $P'(\mathbf{h}; E)$ we denote the corresponding set of sums when both $\mathscr{L}$ and $\mathscr{R}$ contain all intervals. Clearly

$$P(\mathbf{h}; \mathbf{A}; E) \subseteq P'(\mathbf{h}; E) \tag{57.1}$$

If $h_s = H(s = l, r)$, where H is finitely additive, then we have equality in (57.1). For we can obtain from **A** a division of each interval contained in E, so that to each finite family of disjoint intervals of E there corresponds a finite family of disjoint intervals from **A**. By the finite additivity of H the corresponding sums are the same.

When the values of H lie in a normed linear space with norm $\|.\|$, its variation is usually defined to be

$$\sup\|z\|, \quad z \in P'(H; E)$$

and if this is finite, H is said to be of *bounded variation*. A corresponding definition in the linear topological space Z is obtained as follows. A set P *is bounded*, if, given V in $\mathscr{N}$, there is an integer j, depending on P and V, such that $P \subseteq jV$. Further, H *is of bounded variation* if $P'(H; E)$ is bounded. See, for example, Weston (1957). It is clear that we can replace $P'(H; E)$ by the equal set $P(H; \mathbf{A}; E)$, and this opens the way to a generalization useful to us.

We say that a pair $\mathbf{h}$ of interval functions is of *bounded variation in an elementary set E* if, given V in $\mathscr{N}$, there are $\mathbf{A}_1$ complete in E, and an integer $j > 0$, such that

$$P(\mathbf{h}; \mathbf{A}_1; E) \subseteq jV \tag{57.2}$$

Also $\mathbf{h}$ *is of variation zero in E* if in (57.2) we can always take $j = 1$. Two pairs $\mathbf{h}, \mathbf{h}^*$ *are variationally equivalent in E* if $\{h_l - h_l^*, h_r - h_r^*\}$ is of variation zero in E. Definitions of *bounded variation, variation zero, and variational equivalence, in a set X, relative to E*, are obtained on substituting

$$\{ch(X)h_l, ch(X)h_r\}$$

for $\mathbf{h}$. For simplicity we write

$$P(\mathbf{h}; \mathbf{A}; E; X) = P(\{ch(X)h_l, ch(X)h_r\}; \mathbf{A}; E)$$

In Henstock (1960a), p. 47, there is defined a variational integral for functions with values in a normed linear space, and this would correspond to taking a sum of norms for P, instead of a sum of the h_s, and then taking the norm. Thus our present definitions are more general than Henstock (1960a), p. 47, and include it as a special case. On the other hand, our present definition is a special case of the very general definition of Henstock (1960a), p. 48. Here our base space is the real axis, and our space of values is linear; the sets S are intervals open on the right, the family $\mathscr{I}$ is that of all such intervals in an elementary set E, and the family $\mathscr{J}$ is $\mathscr{L} \cup \mathscr{R}$, for $\mathbf{A}$ complete in E. The space $\mathscr{G}$ of neighbourhoods N of a point of $\mathscr{H}$ becomes the family $\mathscr{N}$ of neighbourhoods V of the point θ of Z. The mapping $\mathscr{M}(x, y)$ of $\mathscr{G}$ into $\mathscr{H}$ becomes the mapping of Z into itself that sends z into $z+y-x$. Further, in Henstock (1960a), $\chi(S)$ is a function of sets S whose values are sets of $\mathscr{H}$. It was intended that

$$\chi(S) \subseteq \chi(S') \textit{ if } S \subseteq S' \tag{57.3}$$

Let S, S' be disjoint, with union another member of $\mathscr{I}$. Then we are given in Henstock (1960a) that $\chi(S)$, $\chi(S')$ are disjoint, so that by (57.3),

$$\chi(S) \cup \chi(S') \subseteq \chi(S \cup S') \tag{57.4}$$

corresponding to finite superadditivity in Henstock (1961b). For our sets S, (57.4) is

$$\chi(u, v) \cup \chi(v, w) \subseteq \chi(u,w)(u < v < w) \tag{57.5}$$

The many-valued function $G(S)$ is the two-valued $\mathbf{h}$, while $H(S)$ is the single-valued function $H(u, v)$ of intervals; and $\mathbf{h}$, H are *χ-variationally equivalent in E* if, given V in $\mathscr{N}$, there are $\mathbf{A}_2$, complete in E, and a χ satisfying (57.5), such that

$$\chi(I) \subseteq V(I \subseteq E) \tag{57.6}$$

and that if $[u, v)$ is in $\mathscr{L}_2$, there is an x in $\chi(u, v)$ such that

$$H(u, v) - h_l(u, v) \in \chi(u, v) - x \tag{57.7}$$

and similarly for h_r and $[u, v)$ in $\mathscr{R}_2$. Thus the $\chi(u, v) - x$ are the neighbourhoods of the origin, while the $\chi(u, v)$ themselves are displaced neighbourhoods fitting together to form a set in V.

For example, if Z is the space of reals and if λ is a strictly increasing point function, we could have

$$\chi(u, v) = [\lambda(u), \lambda(v)), x = \tfrac{1}{2}\{\lambda(u) + \lambda(v)\}$$
$$\chi(u, v) - x = [\tfrac{1}{2}\{\lambda(u) - \lambda(v)\}, \tfrac{1}{2}\{\lambda(v) - \lambda(u)\})$$

We now adapt this definition to the requirements of the linear space Z, replacing $\chi(u, v) - x$ by $\chi_1(u, v)$ and changing (57.5) to

$$\chi_1(u, v) + \chi_1(v, w) \subseteq \chi_1(u, w) \textit{ (all } u, v, w \textit{ in } a \leqslant u < v < w \leqslant b) \tag{57.8}$$

Here we call such a χ_1 *finitely superadditive in* $[a, b]$. If also

$$\theta \in \chi_1(u, v) \textit{ if } [u, v) \subseteq [a, b] \tag{57.9}$$

then we easily prove the result corresponding to (57.3), namely

$$\chi_1(u, v) \subseteq \chi_1(u', v') \text{ if } [u, v) \subseteq [u', v') \subseteq [a, b] \qquad (57.10)$$

A function satisfying (57.10) is called *monotone increasing in* $[a, b]$. Replacing (57.7), we say that $\mathbf{h}$ *is of χ_1-bounded variation in* $[a, b]$, if, given V in $\mathscr{N}$, there are $\mathbf{A}_3$ complete in $[a, b]$, a χ_1 satisfying (57.8; 57.9) and an integer $j > 0$, such that

$$h_s(u, v) \in \chi_1(u, v)(s = l,\ [u, v) \in \mathscr{L}_3;\text{ or } s = r, [u, v) \in \mathscr{R}_3) \qquad (57.11)$$

$$\chi_1(a, b) \subseteq jV \qquad (57.12)$$

If we can always take $j = 1$ then $\mathbf{h}$ *is of χ_1-variation zero in* $[a, b]$; and so on.

THEOREM 57.1. *Definitions using* P *and those using* χ_1 *are equivalent.*

Let (57.8; 57.9; 57.11; 57.12) be true, and let $\mathscr{Q}_1$ be a finite set of disjoint intervals from $\mathbf{A}_3$. Then there is another finite set $\mathscr{Q}_2$ of disjoint intervals, disjoint from the intervals of $\mathscr{Q}_1$, such that the intervals of $\mathscr{Q}_1$, $\mathscr{Q}_2$ have union $[a, b)$, with

$$(\mathscr{Q}_1) \sum h_s \in (\mathscr{Q}_1) \sum \chi_1$$
$$\subseteq (\mathscr{Q}_1) \sum \chi_1 + (\mathscr{Q}_2) \sum \chi_1 \subseteq \chi_1(a, b) \subseteq jV$$

Thus (57.2) is true with $\mathbf{A}_1$, E replaced by $\mathbf{A}_3$, $[a, b]$, respectively.

Conversely, let (57.2) be true, with $E = [a, b]$. The function

$$\chi_1(u, v) = P(\mathbf{h}; A_1; [u, v))$$

satisfies (57.12). Taking the empty set for $\mathscr{Q}$, we see that (57.9) is true. Further, if $u < v < w$, a finite collection of disjoint intervals in $[u, v]$, with a similar collection in $[v, w]$, give a similar collection in $[u, w]$. Hence (57.8) is true. Taking a single interval for $\mathscr{Q}$ shows that (57.11) is true with $\mathbf{A}_3$ replaced by $\mathbf{A}_1$. Hence the theorem.

We can easily define the inner variation by obvious modifications.

In previous chapters we have used the numerical value of the variation of our functions. To take its place here we can use $\bar{P}(\mathbf{h}; E)$, the intersection for all $\mathbf{A}_4$ complete in E, of the closures of the sets $P(\mathbf{h}; \mathbf{A}_4; E)$. Using these definitions we can obtain analogues of many previous theorems, defining the variational integral in an obvious way and then showing that it is equivalent to the Riemann-complete integral.

Ex. 57.1. If $\mathbf{h}$ is of variation zero in E, or in X, relative to E, prove that the corresponding P contains the single point θ.

Ex. 57.2. Let X be the union of sets $X_{1j} (j = 1, 2, \ldots)$ Then prove that

$$\bar{P}(\mathbf{h}; E; X) \subseteq \sum_{j=1}^{\infty} \bar{P}(\mathbf{h}; E; X_{1j})$$

Ex. 57.3. Let E_1, E_2 be non-overlapping elementary sets, and let X_2 be a set of points, and $\mathbf{h}$ a pair of interval functions. Then prove that

$$\bar{P}(\mathbf{h}; E_1; X_2) + \bar{P}(\mathbf{h}; E_2; X_2) = \bar{P}(\mathbf{h}; E_1 \cup E_2; X_2)$$

Ex. 57.4. If a, b are scalar constants, $\mathbf{h}, \mathbf{h}^*$ are pairs of interval functions, E an elementary set, X an arbitrary set, prove that

$$\bar{P}(\{ah_l + bh_l^*, ah_r + bh_r^*\}; E; X)$$
$$\subseteq a\bar{P}(\mathbf{h}; E; X) + b\bar{P}(\mathbf{h}^*; E; X)$$

10

INTEGRATION IN STATISTICS

58. Introduction

There are many solutions to the problem of how to apply mathematical probability theory to statistical data. This chapter gives the solution that the author has found most satisfactory, together with a very brief discussion of some other solutions, and an arbitrary selection of one or two topics that illustrate the previous theory.

Applied mathematics, or natural philosophy, is the application of pure mathematics to the natural world. The scientist carries out experiments or observations, and collects data, numerical or otherwise, which he then has to interpret. The mathematician supplies him with various *mathematical models*, and the scientist uses his judgement in selecting the model, and the values of the parameters involved, that will fit the particular situation most closely. The last stage in the process is to make suitable predictions from the mathematical model and to observe whether they are true or false. If the latter, the model is discarded or modified. Nothing, or almost nothing, of this can be exact; exactness can only occur in the realm of pure mathematics. The scientist has to contend with *experimental error* in various forms. For example, in the amplification of varying electric currents there is imposed a random noise produced by the thermal motion of electrons in the amplifier. For some applications a maser operating at very low temperatures can cut out such random noise completely, or almost completely, but there is still the random noise produced between source and amplifier that distorts the signal. Experimental error has as many forms as the number of kinds of experiments performed, and it occurs increasingly as we go from physics to chemistry and biology. In many subjects, such as economics, anthropology, linguistics, one cannot carry out experiments, one can only collect the data as events occur. Here, the possibility of error is greater still. The study of such data, and the errors of experiments, has produced two more branches of applied mathematics, those of *statistics* and *statistical physics*.

In other branches of applied mathematics it is usually vital to smooth or rationalize the results of experiments to obtain a suitable pure mathematical model. If it gives correct predictions we replace the experimental results by the model. This occurs in statistical physics, the pure mathematical model being the theory of *probability*. But in statistics though the model is (roughly) the same, we cannot discard the raw material in favour of it, as this is the source of many fallacies. We have to be careful in keeping separate the raw material and the pure mathematical model, and we are repaid by a resulting greater precision of possible statements in the subject.

59. Classification and relative frequency

The first operation in statistics is that of *classification*. If the data consist of single real numbers x we can divide up their range into a finite number of disjoint sets, for example, the sets given by

$$x \leqslant -n,\ -n < x \leqslant -n+1,\ \ldots,\qquad 0 < x \leqslant 1,\ \ldots,\ n-1 < x \leqslant n,\ x > n$$

to take a very simple case. If the data gives a fixed number p of real numbers we can represent each value as a point of p-dimensional Cartesian space, and we can again divide up the space into a finite number of disjoint sets. In practice we cannot observe a complete infinite sequence of real numbers, but we may have data in which p can be arbitrarily large, in which case we use an infinite dimensional Cartesian space, which can again be divided up in a suitable way into a finite number of disjoint sets, by dividing the space only with respect to a finite number of coordinates and leaving the remainder undivided. We can observe continuous functions, an early example being the traces produced by a barograph. We take a finite number of values of the independent variable and classify the values of the functions at those points. Some variables cannot be measured, or can only be measured at great inconvenience, for example, the grading by eye of colours or shades; here, classification becomes difficult.

Let us suppose that the classification has been carried out, and that we are given a finite number, say, $S_1, S_2, \ldots, S_n$ of disjoint sets, and a finite number of data values. We count the number N_j of times that the data values fall into the set S_j, for $j = 1, 2, \dots, n$. The total number of values is

$$N = \sum_{j=1}^{n} N_j$$

The relative frequency of values in the set S_j is defined to be $f_j{}^N = N_j/N$, *satisfying*

$$0 \leqslant f_j{}^N \leqslant 1 \tag{59.1}$$

$$\sum_{j=1}^{n} f_j{}^N = 1 \tag{59.2}$$

If $j \neq k$, the relative frequency of values in $S_j \cup S_k$ is $f_j{}^N + f_k{}^N$, so that the relative frequency is finitely additive.

If we have made a proper classification, and if we increase the amount of data, it usually happens that the $f_j{}^N$ tend to fixed values. If they do not, we may have neglected an important source of variation. For example, we might think that a particular process is independent of time, but when data are collected, the data values may tend to rise with time.

60. Probability

In the pure mathematical model we attach a number p_j, called the *probability of* S_j, to the set $S_j (j = 1, 2, \dots, n)$. If it is to represent an ideal form of the relative frequency it ought to obey conditions similar to (59.1; 59.2), namely,

$$0 \leqslant p_j \leqslant 1 \tag{60.1}$$

$$\sum_{j=1}^{n} p_j = 1 \tag{60.2}$$

while the probability to be attached to a finite union of the S_j is to be the sum of the corresponding probabilities of the separate S_j, so that the probability is to be *finitely additive*.

The problem now arises of linking the probability and the relative frequency. An obvious way is to write

$$p_j = \lim_{N \to \infty} f_j{}^N \tag{60.3}$$

but this cannot be given a rigid interpretation. First, as has already been pointed out, in practice we cannot observe all terms of an infinite sequence without taking an infinite time over it. Secondly, even if we could, and if the limit exists for one set of data, there is nothing ~~that~~ to ensure~~s~~ that the limit will exist for another set of the same kind of data, and that the two limits are the same, except for the trivial cases when $f_j{}^N = 0$ for all N, or $f_j{}^N = 1$ for all N, and similar cases.

For example, we might classify bank notes according to the last digit of the number printed on them, so that $j = 0, 1, \dots, 9$. Then when notes are collected from the bank, on the one hand, the bank clerk might issue new notes alone, so that, apart from printing errors, the integers j occur in rotation; or, on the other hand, an eccentric bank clerk might pick out those notes having a number with last digit always, say, 4, and issue these, so that $N_j = 0 (j \neq 4)$.

To make this kind of definition amenable to pure mathematics, von Mises defined a *Kollektiv* as an infinite sequence $\{x_j\}$ of 0's and 1's with the properties that the relative frequency of 0's in the first N members of the sequence tends to a limit as N tends to infinity, and that the same is true, with the same limit, for an arbitrary subsequence of the $\{x_j\}$. By taking the subsequence to consist of 1's alone, or of 0's alone, it is clear that the only Kollektivs are those sequences for which either $x_j = 0$ for all j greater than some fixed J, or $x_j = 1$ for all $j > J$. It was not the intention of von Mises to restrict his Kollektivs so drastically.

There is therefore some restriction to be applied to the data of the example, in that the bank notes should be thoroughly mixed or randomized. Again, the subsequence of a Kollektiv was supposed to be chosen without knowing the values x_j. But the accurate definition of such ideas would involve probability in a circular fashion, so that although definition (60.3) is the easiest one to take, it cannot withstand accurate examination.

Let us now examine the Kollektivs in the light of the integration theory. We can write the $\{x_j\}$ as a decimal in the scale of 2, and between 0 and 1, namely,

$$x = 0.x_1 x_2 \dots x_j \dots \tag{60.4}$$

Each number in [0, 1] corresponds to a sequence $\{x_j\}$, and different sequences $\{x_j\}$, $\{x_j'\}$ can give the same x only when there is an integer J such that

$$x_J = 0,\ x_j = 1 (j > J)$$
$$x_J' = 1,\ x_j' = 0\ (j > J)$$

L

or the similar result with $\{x_j\}$, $\{x_j'\}$ interchanged. Such points x are multiples of powers of $\frac{1}{2}$, and the set of those x with a double representation is therefore countable and can be neglected.

Let us take constants m_1, m_2 such that

$$0 < m_j < 1 \ (j = 1, 2), \quad m_1 + m_2 = 1 \tag{60.5}$$

Then there is a continuous non-negative and finitely additive interval function p satisfying the recurrence relations

$$\begin{aligned} &p(0, 1) = 1, \\ &p(2m.2^{-j}, (2m+1)2^{-j}) = m_1 p(m.2^{-j+1}, (m+1)2^{-j+1}), \\ &p((2m+1)2^{-j}, (2m+2)2^{-j}) \\ &\qquad = m_2 p(m.2^{-j+1}, (m+1)2^{-j+1}) \end{aligned} \tag{60.6}$$

If $x_1 + x_2 + \ldots + x_N = M$, the corresponding numbers lying in a set X_{MN}, then M of the first N of the x_j are 1 and the rest are 0. The set X_{MN} is a finite union of intervals. Thus we obtain

$$V(p; [0, 1]; X_{MN}) = {}^N C_M m_1^{N-M} m_2^M = w_M, \textit{ say.} \tag{60.7}$$

Given $\epsilon > 0$, let k be the integer next lower than $Nm_2 + N\epsilon$. Then for $j > k$,

$$\begin{aligned} \frac{w_{j+1}}{w_j} = \frac{(N-j)m_2}{(j+1)m_1} &< \frac{(N - Nm_2 - N\epsilon)m_2}{(Nm_2 + N\epsilon)m_1} \\ &= \frac{1 - \epsilon/m_1}{1 + \epsilon/m_2} = \eta, \ 0 < \eta < 1 \end{aligned}$$

Thus if $X^*_{N\epsilon}$ is the set where $x_1 + x_2 + \ldots + x_N > m_2 + 2\epsilon N$, we use (60.7), obtaining

$$\begin{aligned} V(p; [0, 1]; X^*_{N\epsilon}) &= \sum_{j > m_2 + 2\epsilon N} w_j < \\ &< w_{k+1} \eta^{\epsilon N - 1}/(1-\eta) < \eta^{\epsilon N - 1}/(1-\eta) \\ V\Big(p; [0, 1]; \bigcup_{N \geqslant M} X^*_{N\epsilon}\Big) &< \eta^{\epsilon M - 1}/(1-\eta)^2 \end{aligned} \tag{60.8}$$

Hence the set X_ϵ where

$$\lim_{N \to \infty} \sup(x_1 + x_2 + \ldots + x_N)/N > m_2 + 2\epsilon \textit{ satisfies}$$

$$V(p; [0, 1]; X_\epsilon) = 0 \tag{60.9}$$

Similarly for $-\epsilon$ in place of ϵ, and then by taking

$$\epsilon = 1/n, \ n = 1, 2, \ldots$$

we see that except for a set of x satisfying (60.9), we have

$$\lim_{N \to \infty} (x_1 + x_2 + \ldots + x_N)/N = m_2 \tag{60.10}$$

Thus if we use p as measure, almost all Kollektivs give the limit m_2. Changing m_2, in the region (60.5), we do not change the limit given by a particular Kollektiv, but we change the sets of variation zero. There is no criterion *depending on* $\{x_j\}$ *alone*, to tell us whether the corresponding number x lies in the set satisfying (60.10), or in the complement of that set; the criterion must depend on the particular value of m_2 chosen, m_2 must be chosen first, and cannot be deduced from the behaviour of $\{x_j\}$. This gives a deeper reason why (60.3) must be rejected as a definition of probability.

Another definition of probability has been given by Keynes (1921), Jeffreys (1939), and Ramsey (1931). They said that probability is the *degree of rational belief*, and by using various axioms, are able to build a theory on that foundation. However, the resulting probability is rather subjective, and no method is given for evaluating the numerical values of most probabilities involved in statistical work.

Finally, Kolmogoroff (1933) defines probability as a pure mathematical measure, and develops the pure mathematics of the subject without bringing in the data.

61. Probability in three areas

Each definition in the previous section seems to hold part but not all of the ethos of probability in statistics, so that a blend of all three seems to be indicated. My main assumption is that the numerical work of the applied statistician is correctly carried out, but perhaps wrongly interpreted. Much of the work involves tests of significance, so that we concentrate on these.

When considering a possible statistical experiment we usually assume that a certain effect will not occur, and so we have a *null hypothesis*. Probability here is the degree of our rational belief in the experiment, and to it we assign a number δ in $[0, \frac{1}{2}]$, called the *level of significance*. If the null hypothesis is a basic belief we put $\delta = 0$ and have no need to do the experiment. If, however, it is reasonable to us to perform the experiment, we take δ in $(0, \frac{1}{2}]$, with value such that we accept a situation in which we reject the hypothesis, when true, in a proportion δ of cases. The more costly a rejection is to us, the nearer we take δ to 0. Experiments in parapsychology have the smallest δ. Here we test whether we can obtain information of the natural world or the thoughts of an agent, or whether bodies can be moved, by ways that do not use the known senses or sciences. Many scientists would say that such an effect cannot occur, and would put $\delta = 0$. Others, conceding that the effect might happen, would have to be thoroughly convinced before accepting the effect as operative, and these put $\delta = 10^{-10}$, or some other value still nearer to but greater than 0.

Having chosen δ, we arrange the test procedure so that when the hypothesis is true we reject it in a proportion δ of cases, and accept it in a proportion $1-\delta$ of cases, both proportions being true in the long run. The meaning of the latter phrase is defined mathematically, probability being a finitely additive non-negative set function with greatest attained value 1. The choice of a suitable *probability measure* to fit our particular problem, and the choice of significance level, are together known as the *problem of specification*. The choice is often governed by the knowledge of previous experiments, by some theoretical argument, or by some hypothesis to be tested.

We arrange to obtain a sample of n independent values $y_1, \dots, y_n$ of a variable from the experiment that has not yet begun. The integer n may be fixed or may be a statistical variable itself. Pure mathematics can partially but never wholly solve the problem of ensuring that the values are independent, there is still much in the design of the experiment that has to be left to the individual assessment of the applied statistician. From the probability measure of the variable we find the probability measure of the sample, and we choose a set $C(\delta)$ of samples that has $1-\delta$ for probability measure, where δ is the significance level. If the variable is discrete this is not always possible, and we sometimes have to use sets $C(\delta)$ with probability measure greater than $1-\delta$. When n is variable the $C(\delta)$ depends on n, and we also use $C(\delta')$ for δ' in $(\delta, 1)$ and near to 1, with $C(\delta) \supset C(\delta')$.

If n is fixed, we carry out the experiment and obtain the sample. It is *compatible* with the null hypothesis, and is *not significant*, relative to δ and $C(\delta)$, if the sample lies in $C(\delta)$. Otherwise the sample is *incompatible* with the hypothesis, and is *significant*, and the null hypothesis is rejected. If n is variable, and if the sample is in $C(\delta')$, the sample is *compatible* with the hypothesis and is *not significant*, relative to δ' and $C(\delta')$, and the null hypothesis is accepted. If the sample is not in $C(\delta)$ it is *incompatible* with the hypothesis and is *significant* relative to δ and $C(\delta)$, and the null hypothesis is rejected. If the sample is in $C(\delta)$ but not in $C(\delta')$ we take a further value y_{n+1}, and retest. Here we encounter probability in the third area, that of the data.

62. Independence

Let a variable x range over a set X, and a variable y over Y. Then the set consisting of values of the combined variable (x, y), for all $x \in X$ and all $y \in Y$, is written as $X \times Y$, *the Cartesian product of* X *and* Y.

In this book we suppose that x and y lie on the real line, so that (x, y) is a point of the plane defined by two intersecting lines at right angles. We also suppose that the sets of the simple classification of Section 59 are all intervals, so that the probability measure, the pure mathematical model of the probability of Sections 60 and 61, is a finitely additive interval function. We can suppose that it is defined for all real intervals, say as $p_x(I)$ for x and $p_y(I)$ for y. We suppose also that there is a finitely additive rectangle function $p(R_w-)$ attached to the rectangles R_w- of the plane, that will serve as the probability measure attached to (x, y). If X is a set of x, and if the integral exists, then

$$(V)\int_{-\infty}^{\infty} ch(X)\,\mathrm{d}p_x = V(p_x; (-\infty, \infty); X)$$

is the probability measure $p_x^*(X)$ to be attached to the set X. Note that for the interval $I = [u, v)$,

$$p_x^*((u, v)) \leqslant p_x(u, v) \leqslant p_x^*([u, v])$$

since p_x is non-negative, with equality if p_x is continuous at

u and v. Thus we have no need to define $p_x{}^*$ as a completely additive set function, we need only begin with the finitely additive interval function p_x, and then Theorem 38.1 (p. 89) gives what we require. Similarly we can define $p_y{}^*$, p^* from p_y, p. If for all intervals I, J we have

$$p(I\times J) = p_x(I)p_y(I) \tag{62.1}$$

we say that *x and y are independent*. Strictly, this is not a property of the variables x, y, but of the models p_x, p_y, p, of the three probabilities that we are using. From Theorem 44.2 (p. 109) (Fubini),

$$p^*(X\times Y) = p_y{}^*(X)p_y{}^*(Y) \tag{62.2}$$

Similar results can be obtained for n dimensions.

Returning to the set $C(\delta)$ of Section 61, a set of points $s = (x_1, ..., x_n)$, we are given that

$$p_s{}^*(C(\delta)) = 1-\delta$$

If in N *independent* cases the sample s is in $C(\delta)$ exactly M times, we can compute the probability measure by using finite additivity, (62.2), and the property of combinations as in the binomial theorem, to obtain

$${}^N C_M (1-\delta)^M \delta^{N-M}$$

As in the proof of (60.8; 60.9; 60.10) from (60.7) we see that $M/N \to 1-\delta$ as $N\to\infty$, except for a set of samples of probability measure 0. This gives a pure mathematical definition of the statement that M/N tends to the right value in the long run, and it is the simplest of such convergence results, deeper results also being true.

In books on statistics there are given tests using *contingency tables*, and basically these are tests of the hypothesis that the variables involved are independent.

63. The Neyman–Pearson theory of tests of simple hypotheses

There now arises the problem of finding, in some sense, the best possible region $C(\delta)$. Neyman and Pearson (1933) say that we commit an *error of the first kind* when we reject a hypothesis when it is true; and we commit an *error of the second kind* if we accept the hypothesis when it is false. Their idea is to fix the probability measure of errors of the first kind,

and to minimize the probability measure of errors of the second kind. The two probability measures are not the same, the former being p_0, say and the latter being an alternative probability measure, say p_1. Thus here we can only test one hypothesis, giving rise to p_0, against another hypothesis, giving rise to another single function p_1.

We suppose for simplicity that p_0, p_1 are continuous. As they are non-negative they are clearly AC^* with respect to p_0+p_1, and then Theorem 34.5 gives the existence of $f_j \geqslant 0$ p. such that

$$p_j = (V)\int f_j\,\mathrm{d}(p_0+p_1)\ (j = 0, 1)$$

Theorem 63.1. *If for fixed* $\alpha > 0$, U *is a set of samples* s *satisfying*

$$\mathrm{f}_1(\mathrm{s}) \leqslant \alpha f_0(\mathrm{s}) \tag{63.1}$$

then

$$\mathrm{p}_1{}^*(\mathrm{U}) \leqslant \alpha\mathrm{p}_0{}^*(\mathrm{U}) \tag{63.2}$$

If in U_1,

$$\mathrm{f}_1(s) > \alpha\mathrm{f}_0(\mathrm{s}) \tag{63.3}$$

then either

$$\mathrm{p}_1{}^*(\mathrm{U}_1) > \alpha\mathrm{p}_0{}^*(\mathrm{U}_1) \tag{63.4}$$

or

$$\mathrm{p}_0{}^*(\mathrm{U}_1) = 0 \tag{63.5}$$

Here (63.2) is a simple application of Theorem 30.1 (30.3), p.5 and in (63.4) we obtain $\geqslant$. If equality in (63.4) then

$$(V)\int_{-\infty}^{\infty} ch(U_1)(f_1-\alpha f_0)\,\mathrm{d}(p_0+p_1) = 0$$

with (63.3), so that by Ex. 31.1 p.64 $ch(U_1)(f_1-\alpha f_0) = 0$ a.e., and we have (63.5).

Theorem 63.2. *Let* $\mathrm{U}(\alpha)$ *be the set of* s *satisfying* (63.1). *If a set* U *satisfies*

$$\mathrm{p}_0{}^*(\mathrm{U}) = \mathrm{p}_0{}^*(\mathrm{U}(\alpha)),\quad \mathrm{p}_0{}^*(\mathrm{U}\cap \mathrm{CU}(\alpha)) > 0 \tag{63.6}$$

then

$$p_1{}^*(\mathrm{U}) > \mathrm{p}_1{}^*(\mathrm{U}(\alpha)) \tag{63.7}$$

Sets $U(\alpha)\cap CU$ and $U\cap CU(\alpha)$ are in regions where s

satisfies (63.1; 63.3) respectively, so that by Theorem 63.1,

$$\begin{aligned} p_1^*(U \cap CU(\alpha)) &> \alpha\, p_0^*(U \cap C\, U(\alpha)) \\ &= \alpha p_0^*(U) - \alpha p_0^*(U \cap U(\alpha)) \\ &= \alpha p_0^*(U(\alpha)) - \alpha p_0^*(U \cap U(\alpha)) \\ &= \alpha p_0^*(U(\alpha) \cap CU) \geqslant p_1^*(U(\alpha) \cap CU), \end{aligned}$$

$$\begin{aligned} p_1^*(U) &= p_1^*(U \cap U(\alpha)) + p_1^*(U \cap CU(\alpha)) \\ &> p_1^*(U(\alpha) \cap CU) + p_1^*(U \cap U(\alpha)) = p_1^*(U(\alpha)) \end{aligned}$$

The theorem shows that if $1-p_0^*(U(\alpha))$ is used as a significance level, the set $U(\alpha)$ is the best possible set, in the sense of Neyman and Pearson, to distinguish between the two hypotheses. Sir R. A. Fisher calls f_0 and f_1 *the likelihoods, and* f_1/f_0, *the likelihood ratio*. If f_1 depends on a parameter β, then so does $U(\alpha)$, which may vary as we vary β. Thus the best set for one value of β may be far from the best for another value of β, so that if we are testing p_0 against the composite hypothesis that produces p_1 with different values of β, we have to make a compromise. Fisher suggests that we should maximize f_1, so obtaining β as a function of s, and substitute this in f_1/f_0, using the new ratio to define new $C(\delta)$. But we cannot go any further into these details.

64. Generating functions, moments

Various integrals are associated with a probability measure p, of the form

$$(V)\int_{-\infty}^{\infty} f(x)\mathrm{d}p$$

For example, there are the mean $\mu(f = x)$, variance $\sigma^2(f = (x-\mu)^2)$, the rth moment for positive integers $r(f = x^r)$, the generating function $(f = x^t)$, moment generating function $(f = e^{xt})$, characteristic function $(f = e^{ixt})$, and so on. As these f are continuous and the p are bounded, finitely additive and non-negative, the integrals are equal to the corresponding Radon integrals.

Another integral is obtained in the following way. We suppose that every probability interval function is continuous on the right, so that

$$p(-\infty, t) = p^*((-\infty, t]),$$

and we write the common value as $p(t)$, with

$$p(u, v) = p(-\infty, v) - p(-\infty, u) = p(v) - p(u)$$

Let x, y be independent variables with probability measures p_x, p_y, and let $X(t)$ be the set of (x, y) with $x+y \leqslant t$. Then the probability measure in the range $(-\infty, t]$, to be attached to $x+y$, is

$$(V)\int_P ch(X(t))\,\mathrm{d}p_x p_y$$

where P is the whole plane. By Fubini's Theorem 44.2 p109 this is

$$\begin{aligned} &= \int_{-\infty}^{\infty} \{(V)\int_{-\infty}^{\infty} ch(X(t))\,\mathrm{d}p_x\}\,\mathrm{d}p_y \\ &= (V)\int_{-\infty}^{\infty} p_x^*((-\infty, t-y])\,\mathrm{d}p_y = (V)\int_{-\infty}^{\infty} p_x(t-y)\,\mathrm{d}p_y \end{aligned}$$

This is known as the *convolution* $p_x{*}p_y$ *of* p_x and p_y.

We could now proceed to the Central Limit and allied theorems. But as these are given in many treatises on probability I will instead consider *Karl Pearson's correlation ratio* η_{yx}, which is concerned with the dependence of two variables, say, x and y. Let the probability measure associated with points $w = (x, y)$ of the plane P be obtained from

$$(V)\int f(w)\,\mathrm{d}w$$

We write

$$g(x) = (V)\int_{-\infty}^{\infty} f(x, y)\,\mathrm{d}y,$$

$$\bar{y}(x) = (V)\int_{-\infty}^{\infty} yf(x, y)\,\mathrm{d}y/g(x), \quad \bar{y} = (V)\int_P yf(w)\,\mathrm{d}w$$

Then η_{yx} is given by

$$\begin{aligned} &1-\eta_{yx}^2 \\ &\quad = (V)\int_P \{y-\bar{y}(x)\}^2 f(w)\,\mathrm{d}w/(V)\int_P (y-\bar{y})^2 f(w)\,\mathrm{d}w \qquad (64.1) \end{aligned}$$

We have the following identity, using Ex. 43.5 p.105

$$(V)\int_P \{y-\bar{y}(x)\}^2 f(w)\,\mathrm{d}w$$

$$=(V)\int_P (y-\bar{y})^2 f(w)\,\mathrm{d}w+2(V)\int_P (y-\bar{y})\{\bar{y}-\bar{y}(x)\}f(w)\,\mathrm{d}w+$$

$$+(V)\int_P \{\bar{y}-\bar{y}(x)\}^2 f(w)\,\mathrm{d}w$$

$$=(V)\int_P (y-\bar{y})^2 f(w)\,\mathrm{d}w+$$

$$+2(V)\int_{-\infty}^{\infty} \{\bar{y}-\bar{y}(x)\}\left\{(V)\int_{-\infty}^{\infty} (\bar{y}-\bar{y})f\,\mathrm{d}\bar{y}\right\}\mathrm{d}x+$$

$$+(V)\int_{-\infty}^{\infty} \{\bar{y}-\bar{y}(x)\}^2 g(x)\,\mathrm{d}x$$

$$=(V)\int_P (y-\bar{y})^2 f(w)\,\mathrm{d}w+2(V)\int_{-\infty}^{\infty} \{\bar{y}-\bar{y}(x)\}\{\bar{y}(x)-\bar{y}\}g\,\mathrm{d}x+$$

$$+(V)\int_{-\infty}^{\infty} \{\bar{y}-\bar{y}(x)\}^2 g(x)\,\mathrm{d}x$$

$$(V)\int_P \{y-\bar{y}(x)\}^2 f(w)\,\mathrm{d}w$$

$$=(V)\int_P (y-\bar{y})^2 f(w)\,\mathrm{d}w-(V)\int_{-\infty}^{\infty} \{\bar{y}-\bar{y}(x)\}^2 g(x)\,\mathrm{d}x$$

It follows that

$$\eta_{yx}^2=(V)\int_{-\infty}^{\infty} \{\bar{y}-\bar{y}(x)\}^2 g(x)\,\mathrm{d}x\Big/(V)\int_P (y-\bar{y})^2 f(w)\,\mathrm{d}w \qquad (64.2)$$

THEOREM 64.1. *If* $\eta_{\mathrm{yx}}=\pm 1$ *then* $\mathrm{y}=\bar{\mathrm{y}}(\mathrm{x})$ *except in a set* W *of* $\mathrm{w}=(\mathrm{x},\mathrm{y})$ *with*

$$\mathrm{V}\left((\mathrm{V})\int \mathrm{f}\,d\mathrm{w};\mathrm{P};\mathrm{W}\right)=0 \qquad (64.3)$$

If $\eta_{\mathrm{yx}}=0$ *then* $\bar{\mathrm{y}}(\mathrm{x})=\bar{\mathrm{y}}$ *except in a set* X *of* x *with*

$$\mathrm{V}\left((\mathrm{V})\int \mathrm{g}\,d\mathrm{x};(-\infty,+\infty);\mathrm{X}\right)=0 \qquad (64.4)$$

The results follow from Ex. 31.1, p.64 with (64.1) for (64.3) and (64.2) for (64.4). Thus, apart from the sets of probability ~~which~~ measure zero, there is in the first case a complete functional dependence of y on x, while in the second case the mean of y for each fixed x is independent of the particular x taken.

REFERENCES

Banach, S. (1932). *Théorie des opérations linéares*, Warsaw.

Burkill, J. C. (1932a). The approximately continuous Perron integral, *Math. Z.*, **34**, 270–78.

Burkill, J. C. (1932b). The Cesàro–Perron integral, *Proc. Lond. math. Soc.*, (2)**34**, 314–22.

Burkill, J. C. (1935). The Cesàro–Perron scale of integration, *Proc. Lond. math. Soc.*, (2)**39**, 541–2.

Burkill, J. C. (1936). Fractional orders of integrability, *J. Lond. math. Soc.*, **11**, 220–26.

Burkill, J. C. (1951). Integrals and trigonometric series, *Proc. Lond. math. Soc.*, (3)**1**, 46–57.

Cameron, R. H. and Martin, W. T. (1941). An unsymmetrical Fubini theorem, *Bull. Amer. math. Soc.*, **47**, 121–5.

Getchell, B. C. (1935). On the equivalence of two methods defining Stieltjes integrals, *Bull. Amer. math. Soc.* (2)**41**, 413–8.

Henstock, R. (1946). On interval functions and their integrals, *J. Lond. math. Soc.* **21**, 204–9.

Henstock, R. (1948). On interval functions and their integrals II, *J. Lond. math. Soc.* **23**, 118–23.

Henstock, R. (1960a). A new descriptive definition of the Ward integral, *J. Lond. math. Soc.* **35**, 43–8.

Henstock, R. (1960b). The use of convergence factors in Ward integration, *Proc. Lond. math. Soc.* (3)**10**, 107–21.

Henstock, R. (1960c). The equivalence of generalized forms of the Ward, variational, Denjoy–Stieltjes and Perron–Stieltjes integrals, *Proc. Lond. math. Soc.* (3)**10**, 281–303.

Henstock, R. (1961a). N-variation and N-variational integrals of set functions, *Proc. Lond. math. Soc.* (3)**11**, 109–33.

Henstock, R. (1961b). Definitions of Riemann type of the variational integrals, *Proc. Lond. math. Soc.* (3)**11**, 402–18.

Hildebrandt, T. H. (1926). The Borel theorem, *Bull. Amer. math. Soc.* **32**, 423–74.

Jeffery, R. L. and Miller, D. S. (1945). Convergence factors for generalized integrals, *Duke math. J.* **12**, 127–42.

Jeffery, R. L. (1951). The theory of functions of a real variable, *Math. Expos.* No. 6, Univ. of Toronto Press.

Jeffreys, H. (1939). *Theory of Probability*, Oxford.

Keynes, J. M. (1921). *A Treatise on Probability*, London.

Kolmogoroff, A. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Berlin.

Lusin, N. (1911–12). Sur on théorème fondamentale du calcul intégral (in Russian), *Trans. Moscow math. Soc.* **28**, 270.

von Neumann, J. (1935). On complete topological spaces, *Trans. Amer. math. Soc.* **37**, 1–20.

Neyman, J. and Pearson, E. S. (1933). On the problem of the most efficient tests of statistical hypotheses, *Phil. Trans.* **A231**, 289–337.

Phillips, R. S. (1940). Integration in a convex linear topological space, *Trans Amer. math. Soc.* **47**, 114–45.

Pollard, S. (1923). The Stieltjes integral and its generalizations, *Quart. J. Math.* **49**, 73–138.

Price, G. B. (1940). The theory of integration, *Trans. Amer. math. Soc.* **47**, 1–50.

Ramsey, F. P. (1931). *The Foundations of Mathematics*, pp. 156–211, 256–7, London.

Rickart, C. E. (1942). Integration in a convex linear topological space, *Trans. Amer. math. Soc.* **52**, 498–521.

Robbins, H. (1948). Mixture of two distributions, *Ann. math. Statist.* **19**, 360–69.

Saks, S. (1937). *Theory of the Integral* (2nd English ed.), Warsaw.

Sierpinski, W. (1920). Sur un problème concernant les ensembles measurables superficiellement, *Fundam. Math.* **1**, 112–5.

Sierpinski, W. (1923). Un lemme métrique, *Fundam. Math.* **4**, 201–3.

Taylor, S. J. (1956). An integral of Perron's type defined with the help of trigonometric series, *Quart. J. Math.* (2)**6**, 255–74.

Ward, A. J. (1936). The Perron-Stieltjes integral, *Math. Z.* **41**, 578–604.

Weston, J. D. (1957). Functions of bounded variation in topological vector spaces, *Quart. J. Math.* (2)**8**, 108–11.

Young, W. H. (1910). On a change of order of integration in an improper repeated integral. *Trans. Camb. phil. Soc.* **21**, 361–76.

Young, W. H. and Young, G. C. (1915). On the reduction of sets of intervals, *Proc. Lond. math. Soc.* (2)**14**, 111–30.

Zaanen, A. C. (1953). *Linear Analysis*, Amsterdam.

INDEX

Abel, N. H., 11
Absolute continuity, 58
Absolute integration, 9, 10, and Sections 38, 40
Additive interval function, 30
Additive rectangle function, 103
Additive set function, 32
Almost everywhere (a.e.), 63
Applied mathematics, 1, 148
Approximate Perron integral, 10
Archimedes, 3
Associated point, 21, 100

Banach, S., 11, 127
Beta function, 117
Borel, E., 18, 125
Bounded inner variation, 49
Bounded set in linear topological space, 144
Bounded variation, 46, 47, 144, 146
Burkill, J. C., 2, 6, 8, 10, 11, 26, 32, 34, 105, 121

Calculus, 3–5, 73, 121
  of residues, 115
Cameron, R. H., 110, 111, 116
Cantor, G., 15
Cardinal numbers, 14
Cartesian product, 155
Cauchy, A. L., 122 and Section 46
Central limit, 159
Cesaro–Perron integral, 11
Characteristic function, 158
Chebyshev inequality, 57
Classification, Section 59
Closed interval, 17
Closed operator, 142
Closed rectangle, 100
Closed set, 20
Closure, 18, 20, 129
Compatible, 155
Complement, 14
Complete, 22, 101, 116
Complete additivity, 2, 89, 142, 143, 155, 156
Complete space, 2, 129, 14
Complex integration, 115
Component, 19
Contingency table, 156
Continuous interval functi 32
Continuous operator, 142
Continuous point function,
Convergence, 31, 129
Convergence-factor integra 2, 11, 12
Convergence in mean, 136
Convex topological space,
Convolution, 159
Correlation ratio, 159
Correspondence, 14
Countability, 14

Daniell, P. J., 9
Darboux, J. G., 5, 6
Defining interval, 21
Defining rectangle, 101

Denjoy, A., 2, 10, 11, 19, 98, 121–4 and Section 48
Dense in itself, 20
Density theorem, 80
Derivative (general), 74
Derived set, 20, 71
Disjoint, 14
Distinct, 13
Distributive functional, 27
Distributive operator, 142
Division, 17, 100

Elementary set, 17, 100
Error of first (second) kind, 156
Eudoxus, 3
Everywhere dense, 23, 34
Exhaustions, method of, 3
Experimental error, 148

Family, 13
Fatou, P., 2, 85, 94, 137
Finitely additive function
  intervals, 2, 30, 142, 156
  rectangles, 103
  sets, 32
Finitely subadditive function
  intervals, 30
  rectangles, 103
Finitely superadditive function
  intervals, 30, 145
  rectangles, 103
Fischer, E., 2
Fisher, Sir R. A., 158
Fubini, G., 109, 156, 159
Functional, 27
Fundamental in measure, 139
Fundamental sequence, 31, 129
Fundamental set, 140

Gamma function, 117, 118
Generalized absolute continuity, 58
Generalized bounded inner variation, 56
Generalized bounded variation, 56
Generating function, 158
Getchell, B. C., 2, 8, 11, 25, 41, 121, 142
Greatest lower bound, 16
Greeks, 3

Half-open rectangle, 100
Henstock, R., 2, 3, 8, 10, 11, 22, 100, 101, 124, 144, 145
Hildebrandt, T. H., 124
Hölder, E., 133, 135

Independence, Section 62
Infimum, 16
Inner cover, 48
Inner variation, 48, 49, 146
Inner variation zero, 49
Inside, 17
Intersection, 13
Interval, 17
Interval function, 25
Inverse function, 35
Isolated point, 20

Jacobian, 112, 117
j-complete, 101
Jeffery, R. L., 11, 112
Jeffreys, Sir H., 153
Jessen, B., 112

Keynes, J. M., 153
Kollektivs, 151–3
Kolmogoroff, A., 153

Least upper bound, 16

Lebesgue, H., 2, 3, 9–11, 28, 32, 47, 55, 64, 80, 82, 85, 89, 98, 121, 122, 126, 142
Left-complete, 21
Leibnitz, G. W., 1, 3–5, 71
Likelihood, 158
Lim, lim inf, lim sup (sets), 14
Limit, 3, 140
Limit-point, 20, 129
Linear functional, 2, 11
Linear space, 127
Linear topological space, 2, 11, 128
Lower sum, 5, 6
Lusin, N., 22, 124

Major function, 10, 41, 123
Marcinkiewicz, J., 112
Martin, W. T., 110, 111, 116
Mathematical model, 148
Mean, 158
Mean convergence, Section 54
Measurable, 95
Measure zero, 55
Member, 13
Miller, D. S., 11
Minkowski, H., 134, 135, 139
Minor function, 10, 41, 123
von Mises, R., 151
Moment, 158
Moment generating function, 158
Monotone increasing, 17, 146
Mutually disjoint, 14

Natural philosophy, 148
von Neumann, J., 128
Newton, Sir I., 1, 3–7, 9, 10, 25, 39, 71
Neyman, J., 156, 158
N-integral, 11
Norm, 130
Null function, 139
Null hypothesis, Section
N-variational integral, 1

One-one correspondence
Open interval, 17
Open set, 19
Operator, 3, 142
Orlicz space, 140
Outer measure, 47, 54, 64
Overlapping, 17, 100

Partial sum, 32
Pearson, E. S., 156, 158
Pearson, Karl, 159
Perfect set, 20
Perron, O., 2, 10, 11, 123
Phillips, R. S., 142
Poisson, 11
Pollard, S., 2, 8, 11, 25
121
Price, G. B., 142
Probability, Chapter 10
Probability measure, 153
Product, 24

Radon, J., 2, 9, 11, 12,
122, 158
Ramsey, F. P., 153
Random, 148, 151
Rectangle, 100
Rectangle function, 102
Regular outer measure,
Relative frequency, Sec
59
Rickart, C. E., 142, 143
Riemann, G. F. B., 2, 5
22, 23, 25, 105, 121,
Riemann-complete inte
2, 9, 11, 12, 23, 26,
41, 63, 104, 105, 116,
Riesz, F., 2, 11
Right-complete, 21
Robbins, H., 110, 111, 11

A., 2, 10, 11, 19, 98,
-4 and Section 48
itself, 20
theorem, 80
re (general), 74
set, 20, 71
14
13
tive functional, 27
tive operator, 142
17, 100

ary set, 17, 100
first (second) kind,

, 3
ere dense, 23, 34
ons, method of, 3
ental error, 148

13
., 2, 85, 94, 137
additive function
als, 2, 30, 142, 156
gles, 103
2
subadditive function
als, 30
gles, 103
superadditive func-

als, 30, 145
gles, 103
E., 2
ir R. A., 158
., 109, 156, 159
al, 27
ntal in measure, 139
ntal sequence, 31,

ntal set, 140

unction, 117, 118

Generalized absolute continuity, 58
Generalized bounded inner variation, 56
Generalized bounded variation, 56
Generating function, 158
Getchell, B. C., 2, 8, 11, 25, 41, 121, 142
Greatest lower bound, 16
Greeks, 3

Half-open rectangle, 100
Henstock, R., 2, 3, 8, 10, 11, 22, 100, 101, 124, 144, 145
Hildebrandt, T. H., 124
Hölder, E., 133, 135

Independence, Section 62
Infimum, 16
Inner cover, 48
Inner variation, 48, 49, 146
Inner variation zero, 49
Inside, 17
Intersection, 13
Interval, 17
Interval function, 25
Inverse function, 35
Isolated point, 20

Jacobian, 112, 117
j-complete, 101
Jeffery, R. L., 11, 112
Jeffreys, Sir H., 153
Jessen, B., 112

Keynes, J. M., 153
Kollektivs, 151–3
Kolmogoroff, A., 153

Least upper bound, 16

Lebesgue, H., 2, 3, 9–11, 28, 32, 47, 55, 64, 80, 82, 85, 89, 98, 121, 122, 126, 142
Left-complete, 21
Leibnitz, G. W., 1, 3–5, 71
Likelihood, 158
Lim, lim inf, lim sup (sets), 14
Limit, 3, 140
Limit-point, 20, 129
Linear functional, 2, 11
Linear space, 127
Linear topological space, 2, 11, 128
Lower sum, 5, 6
Lusin, N., 22, 124

Major function, 10, 41, 123
Marcinkiewicz, J., 112
Martin, W. T., 110, 111, 116
Mathematical model, 148
Mean, 158
Mean convergence, Section 54
Measurable, 95
Measure zero, 55
Member, 13
Miller, D. S., 11
Minkowski, H., 134, 135, 139
Minor function, 10, 41, 123
von Mises, R., 151
Moment, 158
Moment generating function, 158
Monotone increasing, 17, 146
Mutually disjoint, 14

Natural philosophy, 148
von Neumann, J., 128
Newton, Sir I., 1, 3–7, 9, 10, 25, 39, 71
Neyman, J., 156, 158
N-integral, 11
Norm, 130
Null function, 139
Null hypothesis, Section 61
N-variational integral, 11

One-one correspondence, 14
Open interval, 17
Open set, 19
Operator, 3, 142
Orlicz space, 140
Outer measure, 47, 54, 64, 121
Overlapping, 17, 100

Partial sum, 32
Pearson, E. S., 156, 158
Pearson, Karl, 159
Perfect set, 20
Perron, O., 2, 10, 11, 123, 124
Phillips, R. S., 142
Poisson, 11
Pollard, S., 2, 8, 11, 25, 41, 121
Price, G. B., 142
Probability, Chapter 10
Probability measure, 153, 154
Product, 24

Radon, J., 2, 9, 11, 12, 121, 122, 158
Ramsey, F. P., 153
Random, 148, 151
Rectangle, 100
Rectangle function, 102
Regular outer measure, 54
Relative frequency, Section 59
Rickart, C. E., 142, 143
Riemann, G. F. B., 2, 5–11, 22, 23, 25, 105, 121, 122
Riemann-complete integral, 2, 9, 11, 12, 23, 26, 39, 41, 63, 104, 105, 116, 141
Riesz, F., 2, 11
Right-complete, 21
Robbins, H., 110, 111, 116

Saks, S., 112, 124
Sample, 154
Scalar, 127
Set, 13
Sierpinski, W., 64, 106, 112
Significance level, Section 61
Similar, 14
Slice, 102
Specification, 24, 101, 154
Statistics, 1, 3, Chapter 10
Statistical physics, 1, 148, 149
Step function, Section 39
Stieltjes, T. J., 2, 6, 8–11, 25, 35, 105, 121–3, 126, 130
Subadditive, 30, 103
Superadditive, 30, 103, 145
Supremum, 16
Symmetrical Cesaro–Perron integral, 11

Taylor, S. J., 11
Trigonometric series, 11

Union, 13
Unsymmetrical Fubini theorem, 111
Upper sum, 5, 6

Variance, 158
Variation, 46, 47, 146
Variational equivalence, 39, 47, 104, 144, 145
Variational integral, 2, 11, 12, 39, 41, 63, 104, 105, 147
Variation set, 32, 47, 143
Variation zero, 47, 144, 146
Vitali, G., 112

Ward, A. J., 2, 10, 11, 123–126
Weston, J. D., 144

Young, W. H., 9, 110, 126, 131
Young, W. H. and G. C., 124

Zaanen, A. C., 140
Zygmund, A., 112